



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 116350

TO: Sheridan Swope
Location: REM-3C70/3A70
Art Unit: 1652
Monday, March 15, 2004

Case Serial Number: 09/966880

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen E01A69
Phone: 571-272-2518

BOB
barbara.obryen@uspto.gov

Search Notes

RUSH

| | | <u>Needed</u> | <u>Done</u> | <u>IS</u> |
|-----------|--------------------------|--------------------------------------|-------------|---|
| SID 7 | comprising | FL | ✓ | Need ✓ |
| 80-676 | consisting oligo | oligo | ✓ | Need ✓ |
| SID 8 | comprising | FL | ✓ | Need ✓ |
| | consisting oligo | oligo | ✓ | ✓ |
| SID 11-15 | consisting oligo | FL | ✓ | Need FL ✓ |
| | | oligo No oligo on SID 15 (poly A) | ✓ | 12-14 encompassed in oligo SID 7 search |
| SID 9 | consisting | FL | ✓ | Need ✓ |
| | U | U | ✓ | ✓ |
| SID 10 | consisting | FL | Need ✓ | Need ✓ |
| | U | | | |
| SID 35 | consisting comprising | FL | ✓ | ✓ |

Seq 15 is 3741 - 5912 of Seq 10
Seq 13 is 2592 - 2862 of Seq 10
Seq 12 is 1065 - 1212 of Seq 10
Seq 14 is 3156 - 3271 of Seq 10

Seq 11 is 1032 - 1118 of Seq 9 =

SH 6130
Seq 9 - interference - 10 hrs
Seq 10, 11 - All NA regular - 11 hrs
" " oligomer 11 hrs

65 hours

2 day

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 12:22:03 ; Search time 194 Seconds

(without alignments)
11329.453 Million cell updates/sec

Title: US-09-966-880A-7_COPY_80_676

Sequence: 1 atggacagcctctgtatgaa.....ctgcactcttggaacttga 597

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 243257 seqs, 184079884 residues

Word size : 20

Total number of hits satisfying chosen parameters: 27

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|-------|--|
| 1 | 597 | 100.0 | 2818 | 9 US-09-966-880A-7 Sequence 7, Appl |
| 2 | 273 | 45.7 | 6564 | 9 US-09-966-880A-10 Sequence 10, Appl |
| 3 | 273 | 45.7 | 11204 | 9 US-09-966-880A-35 Sequence 35, Appl |
| 4 | 271 | 45.4 | 271 | 9 US-09-966-880A-13 Sequence 13, Appl |
| 5 | 148 | 24.8 | 148 | 9 US-09-966-880A-12 Sequence 12, Appl |
| 6 | 78 | 13.1 | 116 | 9 US-09-966-880A-14 Sequence 14, Appl |
| 7 | 54 | 9.0 | 2172 | 9 US-09-966-880A-15 Sequence 15, Appl |
| 8 | 32 | 5.4 | 2440 | 9 US-09-966-880A-1 Sequence 1, Appl |
| 9 | 30 | 5.0 | 30 | 9 US-09-966-880A-18 Sequence 18, Appl |
| 10 | 30 | 5.0 | 30 | 9 US-09-966-880A-24 Sequence 24, Appl |
| 11 | 28 | 4.7 | 28 | 9 US-09-966-880A-25 Sequence 25, Appl |
| 12 | 28 | 4.7 | 28 | 9 US-09-966-880A-28 Sequence 28, Appl |
| 13 | 27 | 4.5 | 27 | 9 US-09-966-880A-19 Sequence 19, Appl |
| 14 | 22 | 3.7 | 371 | 10 US-09-918-995-30237 Sequence 30237, A |
| 15 | 22 | 3.7 | 553 | 9 US-09-796-692-4504 Sequence 4504, Ap |

| | | | | |
|------|----|-----|-------|--|
| 16 | 22 | 3.7 | 553 | 14 US-10-040-862-4504 Sequence 4504, Ap |
| 17 | 22 | 3.7 | 553 | 15 US-10-057-475B-4504 Sequence 4504, Ap |
| 18 | 22 | 3.7 | 553 | 15 US-10-154-884B-4504 Sequence 4504, Ap |
| c 19 | 22 | 3.7 | 556 | 9 US-09-796-692-5061 Sequence 5061, Ap |
| c 20 | 22 | 3.7 | 556 | 14 US-10-040-862-5061 Sequence 5061, Ap |
| c 21 | 22 | 3.7 | 556 | 15 US-10-057-475B-5061 Sequence 5061, Ap |
| c 22 | 22 | 3.7 | 556 | 15 US-10-154-884B-5061 Sequence 1, Appl |
| 23 | 22 | 3.7 | 1155 | 15 US-10-460-923-1 Sequence 173, Appl |
| 24 | 22 | 3.7 | 1534 | 9 US-09-729-674-173 Sequence 359, Appl |
| 25 | 22 | 3.7 | 2151 | 9 US-09-822-830A-359 Sequence 26, Appl |
| 26 | 21 | 3.5 | 21 | 9 US-09-966-880A-26 Sequence 1, Appl |
| 27 | 20 | 3.4 | 63158 | 14 US-10-292-198-1 |

ALIGNMENTS

```
RESULT 1
US-09-966-880A-7
Sequence 7, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Muramatsu, Masamichi
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966, 880A
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 2818
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (80)...(673)
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)...(79)
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (677)...(2818)
US-09-966-880A-7
```

Query Match 100.0%; Score 597; DB 9; Length 2818;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| QY | 1 | ATGACAGCCCTTGATGAGAACCGAGAGAGTTTCTTACCAATTCAGAAATGTCGGTGG | 60 |
| DB | 80 | ATGACAGCCCTTGATGAGAACCGAGAGAGTTTCTTACCAATTCAGAAATGTCGGTGG | 139 |
| QY | 61 | GCTAAGGTCGGCTGAGACCTTACTGTGCTAGTGAAGAGCGCTGACAGTGCTACA | 120 |
| DB | 140 | GCTAAGGTCGGCTGAGACCTTACTGTGCTAGTGAAGAGCGCTGACAGTGCTACA | 199 |
| QY | 121 | TCCTTTCACTGACCTTGGTTATCTTGCAGATAGAACGCTGCGACGTGGAATTGCTC | 180 |
| DB | 200 | TCCTTTCACTGACCTTGGTTATCTTGCAGATAGAACGCTGCGACGTGGAATTGCTC | 259 |
| QY | 181 | TTCTCTCGCTACATCTGGAAGTGGACCTTGAACCTTGGCTGCTACCGCGTCACTGG | 240 |
| DB | 260 | TTCTCTCGCTACATCTGGAAGTGGACCTTGAACCTTGGCTGCTACCGCGTCACTGG | 319 |

QY 24 TTACACCTCGAGAGCCCGCTGACACAGCTGCGCCACATGAGGCGCACTTTGTCGAGG 300

Db 320 TTACCTCTCGAGCCCCCTGCTACACAGCTGCGCCACATGAGGCGCACTTTGTCGAGG 379

QY 301 AACCCCACTCAGTCTGAGATCTTCAACCGCGCCCTCTACTTTCTGTGAGAACCGCAAG 360

Db 380 AACCCCACTCAGTCTGAGATCTTCAACCGCGCCCTCTACTTTCTGTGAGAACCGCAAG 439

QY 361 GCTGAGCCCGAGGGGCTGCGCGGCTGCAACGCGCCGGGTGCAATGACCATATGACC 420

Db 440 GCTGAGCCCGAGGGGCTGCGCGGCTGCAACGCGCCGGGTGCAATGACCATATGACC 499

QY 421 TTCAAAGATTATTTTACTGCTGAGAACTCTTTGTGAAAAACAATGAAAGACTTTGAAA 480

Db 500 TTCAAAGATTATTTTACTGCTGAGAACTCTTTGTGAAAAACAATGAAAGACTTTGAAA 559

QY 481 GCCCGGGAAGGGCTGACATGAAAAATTAGTTCGTCTCGACAGACGCTCGGCGATCCTT 540

Db 560 GCCCGGGAAGGGCTGACATGAAAAATTAGTTCGTCTCGACAGACGCTCGGCGATCCTT 619

QY 541 TTGCGCCCTGATAGGTTGATGACTTACAGACGCAATTCGATCTTTGGACCTTTGA 597

Db 620 TTGCGCCCTGATAGGTTGATGACTTACAGACGCAATTCGATCTTTGGACCTTTGA 676

```

RESULT 2
US-09-966-880A-10
; Sequence 10, Application US/09966680A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966-880A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/11918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 6564
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-966-880A-10

```

| | | | | |
|-----------------------|----------------|--------------------|----------|-------------|
| Query Match | 45.7% | Score 273 | DB 9 | Length 6564 |
| Best Local Similarity | 100.0% | Pred. No. 2.7e-139 | | |
| Matches 273 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |

| QY | 155 | AGAAAGGCGTCCACCGTGAATTCCTTCCTCCGCTATCATCTCGACCTGGACCTGAC | 214 |
|----|------|---|-------|
| Db | 2559 | AGAAAGGCGTCCACCGTGAATTCCTTCCTCCGCTATCATCTCGACCTGGACCTGAC | 26458 |
| QY | 215 | CTGGCCGCGTCTACCGCGTCACTCGTGGTCACTCTTGGAGCCCTTGCTACGACTGTGCC | 274 |
| Db | 2650 | CTGGCCGCGTCTACCGCGTCACTCGTGGTCACTCTTGGAGCCCTTGCTACGACTGTGCC | 27099 |
| QY | 275 | GACATGTGGCCGACCTTCTGCGAGGGGAAACCCACCTCACTGCTGAGATCTTCAACCGGC | 334 |
| Db | 2710 | GACATGTGGCCGACCTTCTGCGAGGGGAAACCCACCTCACTGCTGAGATCTTCAACCGGC | 2765 |
| QY | 335 | GCCTCTACTTCTGTGAGGACCGCAAAGGCTGAACCCGAGGGGCTGCGCGCGCTGCACCGCG | 394 |
| Db | 2770 | GCCTCTACTTCTGTGAGGACCGCAAAGGCTGAACCCGAGGGGCTGCGCGCGCTGCACCGCG | 2828 |
| QY | 395 | CCGGGGTCGAATAGCCATCATGACCTTCAAAG | 427 |

Db 2830 CCGGGTCAATAGCCATCATGACCTTCAAAG 2862

```

RESULT 3
US-09-966-880A-35
/ Sequence 35, Application US/09966880A
/ Patent No. US2002016473A1
/ GENERAL INFORMATION:
/ APPLICANT: Muramatsu, Masamichi
/ TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
/ FILE REFERENCE: 06501-088001
/ CURRENT APPLICATION NUMBER: US/09/966, 880A
/ CURRENT FILING DATE: 2001-09-28
/ PRIOR APPLICATION NUMBER: PCT/JP00/01918
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: JP 11-371382
/ PRIOR FILING DATE: 1999-12-27
/ PRIOR APPLICATION NUMBER: JP 11-178999
/ PRIOR FILING DATE: 1999-06-24
/ PRIOR APPLICATION NUMBER: JP 11-87192
/ PRIOR FILING DATE: 1999-03-29
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 35
/ LENGTH: 11204
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-966-880A-35

```

| Query Match | Similarity | 45.7% | Score 273 | DB 9 | Length 11204 |
|-------------|-------------------|--|-----------|--------|--------------|
| Best Local | Similarity 100.0% | Pred. No. 2,8e-139 | | | |
| Matches 273 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 | |
| QY | 155 | AGAACGGCTGCGCACGTGAAATTGCTCTTCTCCGCTACATCTCGACTGGACCTTAGACC | 214 | | |
| Db | 7805 | AGAAAGGCTGCCACACGTGAAATTGCTCTCTCTCCGCTACATCTCGGACTGGAGCTTAGACC | 7864 | | |
| QY | 215 | CTGGCGGCTGCTACCGCGCTCACTGGTTCACTCTCTGGAGGCCCTGTCTACGACTGTGCC | 274 | | |
| Db | 7865 | CTGGCGGCTGCTACCGCGCTCACTGGTTCACTCTCTGGAGGCCCTGTCTACGACTGTGCC | 7922 | | |
| QY | 215 | GACATGTGGCGGACCTTCTTGGAGAGGAACCCCAACCTCAGTCTGAGATCTTCAACGAGGC | 334 | | |
| Db | 7925 | GACATGTGGCGGACCTTCTTGGAGAGGAACCCCAACCTCAGTCTGAGATCTTCAACGAGGC | 7988 | | |
| QY | 335 | GGCTTACTCTGTGTAGGACCGGCAAGGTGAGCCCGAGGGGCTGGGGCGGCTGCACCGCG | 394 | | |
| Db | 7985 | GGCTTACTCTGTGTAGGACCGGCAAGGTGAGCCCGAGGGGCTGGGGCGGCTGCACCGCG | 8044 | | |
| QY | 395 | CCGGGGTGCAAATAGCCATCATGACCTTCAAG | 427 | | |
| Db | 8045 | CCGGGGTGCAAATAGCCATCATGACCTTCAAG | 8077 | | |

```

US-09-966-880A-13
; Sequence 13, Application US/09966880A
; Patent No. US20020164743A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Honjo, Yasuko
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; CURRENT FILING DATE: 2001-09-28
;
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178399
; PRIOR FILING DATE: 1999-06-24
;

```

;; PRIOR APPLICATION NUMBER: JP 11-87192
;; PRIOR FILING DATE: 1999-03-29
;; NUMBER OF SEQ ID NOS: 36
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 13
;; LENGTH: 271
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-966-880A-13

Query Match 45.4%; Score 271; DB 9; Length 271;
Best Local Similarity 100.0%; Pred. No. 3.2e-158;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 AACGGCTGCACCGGAAATGCTCTCCGCTACATCTGGAGCTGGAGCTGACCT 216
DB 1 AAGGGCTGCACGAGTGAATGCTCTTCTCCGCTACATCTGGAGCTGGAGCTGACCT 60
QY 217 GGCCCGCTGACCGGCTACCTGTTCACTCTCTGAGAGCCCTGCTACGACTGACCGCA 276
DB 61 GGCCCGCTGACCGGCTACCTGTTCACTCTCTGAGAGCCCTGCTACGACTGACCGCA 120
QY 277 CATGTCGCGCACTTCTGAGAGGAAACCCCACTCACTGAGAGTCTTACCGCGCGC 336
DB 121 CATGTCGCGCACTTCTGAGAGGAAACCCCACTCACTGAGAGTCTTACCGCGCGC 180
QY 337 CTCTACTTCTGTGAGAGCGCAAGGCTGAGCCGAGGGGCTGCGGCTGACCGCGCC 396
DB 181 CTCTACTTCTGTGAGAGCGCAAGGCTGAGCCGAGGGGCTGCGGCTGACCGCGCC 240
QY 397 GGGGTGCAAAATGACATCATGACCTTCAAG 427
DB 241 GGGGTGCAAAATGACATCATGACCTTCAAG 271

RESULT 5

US-09-966-880A-12
;; Sequence 12, Application US/09966880A
;; Patent No. US20020164743A1
;; GENERAL INFORMATION:
;; APPLICANT: Honjo, Tasuku
;; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
;; FILE REFERENCE: 06501-088001
;; CURRENT APPLICATION NUMBER: US/09/966, 880A
;; PRIOR APPLICATION NUMBER: PCT/JP00/01918
;; PRIOR FILING DATE: 2000-03-28
;; PRIOR APPLICATION NUMBER: JP 11-371382
;; PRIOR FILING DATE: 1999-12-27
;; PRIOR APPLICATION NUMBER: JP 11-178999
;; PRIOR FILING DATE: 1999-06-24
;; PRIOR APPLICATION NUMBER: JP 11-87192
;; PRIOR FILING DATE: 1999-03-29
;; NUMBER OF SEQ ID NOS: 36
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 12
;; LENGTH: 148
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-966-880A-12

Query Match 24.8%; Score 148; DB 9; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.2e-70;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 CCTCTGATGAACCGAGAAATTTCTTACCAATCAAAAATGTCGCTGAGCTAAAGG 68
DB 1 CCTCTGATGAACCGAGAAATTTCTTACCAATCAAAAATGTCGCTGAGCTAAAGG 60
QY 69 TCGGCGTGAAGACTTCTGTGCTAGTGAAGAGGGCTGACAGTCTTACATCTTTTC 128
DB 61 TCGGCGTGAAGACTTCTGTGCTAGTGAAGAGGGCTGACAGTCTTACATCTTTTC 120

QY 129 ACTGACCTTGTGTTATCTTGGCAATAG 156
DB 121 ACTGACCTTGTGTTATCTTGGCAATAG 148

RESULT 6

US-09-966-880A-14
;; Sequence 14, Application US/09966880A
;; Patent No. US20020164743A1
;; GENERAL INFORMATION:
;; APPLICANT: Honjo, Tasuku
;; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
;; FILE REFERENCE: 06501-088001
;; CURRENT APPLICATION NUMBER: US/09/966, 880A
;; PRIOR APPLICATION NUMBER: PCT/JP00/01918
;; PRIOR FILING DATE: 2001-09-28
;; PRIOR APPLICATION NUMBER: PCT/JP00/01918
;; PRIOR FILING DATE: 2000-03-28
;; PRIOR APPLICATION NUMBER: JP 11-371382
;; PRIOR FILING DATE: 1999-12-27
;; PRIOR APPLICATION NUMBER: JP 11-178999
;; PRIOR FILING DATE: 1999-06-24
;; PRIOR APPLICATION NUMBER: JP 11-87192
;; PRIOR FILING DATE: 1999-03-29
;; NUMBER OF SEQ ID NOS: 36
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 14
;; LENGTH: 116
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-966-880A-14

Query Match 13.1%; Score 78; DB 9; Length 116;
Best Local Similarity 100.0%; Pred. No. 3.7e-32;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 466 GAAAGACTTTCAAACCTGGAGAGGCTGCATGAAAATTCAGTCTTCTCCAGACG 525
DB 39 GAAAGACTTTCAAACCTGGAGAGGCTGCATGAAAATTCAGTCTTCTCCAGACG 98
QY 526 CTTCGGCGCATCTCTTTG 543
DB 99 CTTCGGCGCATCTCTTTG 116

RESULT 7

US-09-966-880A-15
;; Sequence 15, Application US/09966880A
;; Patent No. US20020164743A1
;; GENERAL INFORMATION:
;; APPLICANT: Honjo, Tasuku
;; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
;; FILE REFERENCE: 06501-088001
;; CURRENT APPLICATION NUMBER: US/09/966, 880A
;; PRIOR APPLICATION NUMBER: PCT/JP00/01918
;; PRIOR FILING DATE: 2001-09-28
;; PRIOR APPLICATION NUMBER: PCT/JP00/01918
;; PRIOR FILING DATE: 2000-03-28
;; PRIOR APPLICATION NUMBER: JP 11-371382
;; PRIOR FILING DATE: 1999-12-27
;; PRIOR APPLICATION NUMBER: JP 11-178999
;; PRIOR FILING DATE: 1999-06-24
;; PRIOR APPLICATION NUMBER: JP 11-87192
;; PRIOR FILING DATE: 1999-03-29
;; NUMBER OF SEQ ID NOS: 36
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 15
;; LENGTH: 2172
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-966-880A-15

Query Match 9.0%; Score 54; DB 9; Length 2172;
Best Local Similarity 100.0%; Pred. No. 6.1e-19; Indels 0; Gaps 0;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 544 CCCCCTATGAGCTTGATGACTTACGAGACGACATTGCTTGGGACTTTGA 597
Db 1 CCCCCTATGAGCTTGATGACTTACGAGACGACATTGCTTGGGACTTTGA 54

RESULT 8
US-09-966-880A-1
; Sequence 1, Application US/09966880A
; Patent No. US20020164743A1

; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2440
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (93)...(686)
; NAME/KEY: 5' UTR
; LOCATION: (1)...(92)
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: (690)...(2440)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(2440)
; OTHER INFORMATION: n = A, T, C or G
US-09-966-880A-1

Query Match 5.4%; Score 32; DB 9; Length 2440;
Best Local Similarity 100.0%; Pred. No. 7.5e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 226 TACCGGCTACCTGCTTACCTCTGAGGCC 257
Db 318 TACCGGCTACCTGCTTACCTCTGAGGCC 349

RESULT 9
US-09-966-880A-18
; Sequence 18, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27

; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially synthesized primer sequence, 22
US-09-966-880A-18

Query Match 5.0%; Score 30; DB 9; Length 30;
Best Local Similarity 100.0%; Pred. No. 8.5e-06;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 94 GTAGTGAAGAGCGCTGACAGTCTACATCC 123
Db 1 GTAGTGAAGAGCGCTGACAGTCTACATCC 30

RESULT 10
US-09-966-880A-24
; Sequence 24, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially synthesized primer sequence, p14
US-09-966-880A-24

Query Match 5.0%; Score 30; DB 9; Length 30;
Best Local Similarity 100.0%; Pred. No. 8.5e-06;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 AGCCTTGTGATGAACCGAGAGATTCTT 36
Db 1 AGCCTTGTGATGAACCGAGAGATTCTT 30

RESULT 11
US-09-966-880A-25/c
; Sequence 25, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918

;; PRIOR FILING DATE: 2000-03-28
;; PRIOR APPLICATION NUMBER: JP 11-371382
;; PRIOR FILING DATE: 1999-12-27
;; PRIOR APPLICATION NUMBER: JP 11-178999
;; PRIOR FILING DATE: 1999-06-24
;; PRIOR APPLICATION NUMBER: JP 11-87192
;; PRIOR FILING DATE: 1999-03-29
;; NUMBER OF SEQ ID NOS: 36
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 25
;; LENGTH: 28
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Artificially synthesized primer sequence, p16
US-09-966-880A-25

Query Match 4.7%; Score 28; DB 9; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 CACTGACCTTGGTATCTTCCCAATA 155
DB 28 CACTGACCTTGGTATCTTCCCAATA 1

RESULT 12
US-09-966-880A-28/c

;; Sequence 28, Application US/09966880A
;; Patent No. US20020164743A1
;; GENERAL INFORMATION:
;; APPLICANT: Honjo, Tasuku
;; APPLICANT: Muramatsu, Masamichi
;; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
;; FILE REFERENCE: 06501-088001
;; CURRENT APPLICATION NUMBER: US/09/966/880A
;; CURRENT FILING DATE: 2001-09-28
;; PRIOR APPLICATION NUMBER: PCT/JP00/01918
;; PRIOR FILING DATE: 2000-03-28
;; PRIOR APPLICATION NUMBER: JP 11-371382
;; PRIOR FILING DATE: 1999-12-27
;; PRIOR APPLICATION NUMBER: JP 11-178999
;; PRIOR FILING DATE: 1999-06-24
;; PRIOR APPLICATION NUMBER: JP 11-87192
;; PRIOR FILING DATE: 1999-03-29
;; NUMBER OF SEQ ID NOS: 36
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 28
;; LENGTH: 28
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Artificially synthesized primer sequence, p26
US-09-966-880A-28

Query Match 4.7%; Score 28; DB 9; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 516 CTCGACAGCTTCGGCGCATCTTTG 543
DB 28 CTCGACAGCTTCGGCGCATCTTTG 1

RESULT 13
US-09-966-880A-19/c
;; Sequence 19, Application US/09966880A
;; Patent No. US20020164743A1
;; GENERAL INFORMATION:
;; APPLICANT: Honjo, Tasuku
;; APPLICANT: Muramatsu, Masamichi
;; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
;; FILE REFERENCE: 06501-088001

;; CURRENT APPLICATION NUMBER: US/09/966/880A
;; CURRENT FILING DATE: 2001-09-28
;; PRIOR APPLICATION NUMBER: PCT/JP00/01918
;; PRIOR FILING DATE: 2000-03-28
;; PRIOR APPLICATION NUMBER: JP 11-371382
;; PRIOR FILING DATE: 1999-12-27
;; PRIOR APPLICATION NUMBER: JP 11-178999
;; PRIOR FILING DATE: 1999-06-24
;; PRIOR APPLICATION NUMBER: JP 11-87192
;; PRIOR FILING DATE: 1999-03-29
;; NUMBER OF SEQ ID NOS: 36
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 19
;; LENGTH: 27
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Artificially synthesized primer sequence, 25
US-09-966-880A-19

Query Match 4.5%; Score 27; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 277 CATGTGCGGACCTTCTGCGAGGAGAC 303
DB 27 CATGTGCGGACCTTCTGCGAGGAGAC 1

RESULT 14
US-09-918-995-30237

;; Sequence 30237, Application US/09918995
;; Publication No. US20030073623A1
;; GENERAL INFORMATION:
;; APPLICANT: Hysq, Inc.
;; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
;; FILE REFERENCE: 20411-756
;; CURRENT APPLICATION NUMBER: US/09/918/995
;; CURRENT FILING DATE: 2001-07-30
;; PRIOR APPLICATION NUMBER: US/09/235/076
;; PRIOR FILING DATE: 1999-01-20
;; NUMBER OF SEQ ID NOS: 38054
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 30237
;; LENGTH: 371
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-918-995-30237

Query Match 3.7%; Score 22; DB 10; Length 371;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 TTCACTCTCTGAGGCCCCCTGCT 262
DB 254 TTCACTCTCTGAGGCCCCCTGCT 275

20NTS

RESULT 15
US-09-796-692-4504
;; Sequence 4504, Application US/09796692
;; Publication No. US20020198362A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Maimon, Jane
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERA
;; FILE REFERENCE: 2077.001200
;; CURRENT APPLICATION NUMBER: US/09/796/692
;; CURRENT FILING DATE: 2001-03-01
;; PRIOR APPLICATION NUMBER: 60/186,126

```

; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4504
; LENGTH: 553
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: (107)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (153)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (184)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (193)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (325)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (327)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (473)
; OTHER INFORMATION: n=A,T,C or G
; US-09-796-692-4504
```

```

Query Match      3.7%; Score 22; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      241 TTCACTCTGGAGCCCTGCT 262
Db      531 TTCACTCTGGAGCCCTGCT 552
```

```

RESULT 16
US-10-040-862-4504
; Sequence 4504, Application US/10040862
; PUBLICATION NO. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
```

```

; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-01352005
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4504
; LENGTH: 553
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: (107)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (153)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (184)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (193)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (325)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (327)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (473)
; OTHER INFORMATION: n=A,T,C or G
; US-10-040-862-4504
```

```

Query Match      3.7%; Score 22; DB 14; Length 553;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      241 TTCACTCTGGAGCCCTGCT 262
Db      531 TTCACTCTGGAGCCCTGCT 552
```

RESULT 17
US-10-057-475B-4504
Sequence 4504, Application US/10057475B
Publication No. US20040002068A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Clapper, Jonathan David
APPLICANT: Wang, Aijun
APPLICANT: Ordonez, Nadia
APPLICANT: Carter, Lauren
APPLICANT: McNeill, Patricia Dianne
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-014402US
CURRENT APPLICATION NUMBER: US/10/057,475B
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 10979
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4504
LENGTH: 553
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(553)
OTHER INFORMATION: n = G, A, C or T
US-10-057-475B-4504

Query Match 3.7%; Score 22; DB 15; Length 553;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 TTACCTCCTGGAGCCCTGCT 262
DB 531 TTACCTCCTGGAGCCCTGCT 552

RESULT 18
US-10-154-884B-4504
Sequence 4504, Application US/10154884B
Publication No. US20040005561A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc W.
APPLICANT: Corixa Corporation

TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-013521US
CURRENT APPLICATION NUMBER: US/10/154,884B
CURRENT FILING DATE: 2002-05-23
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 11290
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4504
LENGTH: 553
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(553)
OTHER INFORMATION: n = G, A, C or T
US-10-154-884B-4504

Query Match 3.7%; Score 22; DB 15; Length 553;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 TTACCTCCTGGAGCCCTGCT 262
DB 531 TTACCTCCTGGAGCCCTGCT 552

RESULT 19
US-09-796-692-5061/C
Sequence 5061, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERA
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084


```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-5061
;
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5061
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-5061
;
Query Match          3.7%; Score 22; DB 9; Length 556;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
Qy      241 TTACCTCCTGGAGCCCTGCT 262
Db      26  TTACCTCCTGGAGCCCTGCT 5
;
RESULT 20
US-10-040-862-5061/c
; Sequence 5061, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5061
; LENGTH: 556
;

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-5061
;
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5061
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-5061
;
Query Match          3.7%; Score 22; DB 15; Length 556;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
Qy      241 TTACCTCCTGGAGCCCTGCT 262
Db      26  TTACCTCCTGGAGCCCTGCT 5
;
RESULT 21
US-10-057-475B-5061/c
; Sequence 5061, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aljun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Diane
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5061
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-5061/c
; Sequence 5061, Application US/10154884B
;

```

```
/ Publication No. US20040005561A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013521US
/ CURRENT APPLICATION NUMBER: US/10/154,884B
/ CURRENT FILING DATE: 2002-05-23
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 11290
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 5061
/ LENGTH: 556
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-154-884B-5061

Query Match          3.7% Score 22; DB 15; Length 556;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      241 TTCACCTCCTGGAGCCCTGCT 262
Db      26 TTCACCTCCTGGAGCCCTGCT 5

RESULT 23
US-10-460-923-1
/ Sequence 1, Application US/10460923
/ Publication No. US20040009951A1
/ GENERAL INFORMATION:
/ APPLICANT: MALIM, Michael H.
/ APPLICANT: SHEEHY, Ann M.
/ APPLICANT: HARRIS, Reuben S.
/ APPLICANT: BISHOP, Kate N.
/ APPLICANT: NEUBERGER, Michael S.
/ APPLICANT: GADDIS, Nathan C.
/ APPLICANT: SIMON, James H.M.
/ TITLE OF INVENTION: DNA Deamination Mediates Innate Immunity to Retroviral Infection
/ FILE REFERENCE: 22253-74380
/ CURRENT APPLICATION NUMBER: US/10/460,923
/ CURRENT FILING DATE: 2003-06-13
/ PRIOR APPLICATION NUMBER: US 60/388,513
/ PRIOR FILING DATE: 2002-06-13
/ PRIOR APPLICATION NUMBER: US 60/472,952
/ PRIOR FILING DATE: 2003-05-23
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
```

```
/ LENGTH: 1155
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-460-923-1

Query Match          3.7% Score 22; DB 15; Length 1155;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      241 TTCACCTCCTGGAGCCCTGCT 262
Db      844 TTCACCTCCTGGAGCCCTGCT 865

RESULT 24
US-09-729-674-173
/ Sequence 173, Application US/09729674
/ Patent No. US2001003935A1
/ GENERAL INFORMATION:
/ APPLICANT: Jacobs, Kenneth
/ APPLICANT: McCoy, John M.
/ APPLICANT: Lavallie, Edward R.
/ APPLICANT: Collins-Racie, Lisa A.
/ APPLICANT: Evans, Cheryl
/ APPLICANT: Merberg, David
/ APPLICANT: Treacy, Maurice
/ APPLICANT: Agostino, Michael J.
/ APPLICANT: Steinhilber II, Robert J.
/ APPLICANT: Spaulding, Vikki
/ APPLICANT: Wong, Gordon G.
/ APPLICANT: Clark, Hilary
/ APPLICANT: Rechtel, Kim
/ APPLICANT: Genetics Institute, Inc.
/ TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
/ FILE REFERENCE: 6055-64X
/ CURRENT APPLICATION NUMBER: US/09/729,674
/ CURRENT FILING DATE: 2000-12-04
/ PRIOR APPLICATION NUMBER: 09/539,330
/ PRIOR FILING DATE: 2000-03-30
/ NUMBER OF SEQ ID NOS: 283
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 173
/ LENGTH: 1534
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-729-674-173

Query Match          3.7% Score 22; DB 9; Length 1534;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      241 TTCACCTCCTGGAGCCCTGCT 262
Db      894 TTCACCTCCTGGAGCCCTGCT 915

RESULT 25
US-09-822-830A-359
/ Sequence 359, Application US/09822830A
/ Patent No. US20020142952A1
/ GENERAL INFORMATION:
/ APPLICANT: Genetics Institute, Inc.
/ APPLICANT: Wong, Gordon G.
/ APPLICANT: Clark, Hilary
/ APPLICANT: Rechtel, Kim
/ APPLICANT: Agostino, Michael J.
/ APPLICANT: Howes, Steven H.
/ APPLICANT: Resnick, Richard J.
/ APPLICANT: Gulikota, Kamalakr
/ APPLICANT: Graham, James R.
/ TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
/ FILE REFERENCE: GIN 6402
/ CURRENT APPLICATION NUMBER: US/09/822,830A
```

Mon Mar 15 09:28:00 2004

us-09-966-880a-7_copy_80_676.olg20.rmpb

Page 10

```
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,604
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 359
; LENGTH: 2151
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-830A-359
```

```
Query Match          3.7%; Score 22; DB 9; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      241 TTCACCTCCTGGAGCCCTGCT 262
          |||||
DB      1533 TTCACCTCCTGGAGCCCTGCT 1554
```

```
RESULT 26
US-09-966-880A-26
; Sequence 26, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially synthesized primer sequence, p17
US-09-966-880A-26
```

```
Query Match          3.5%; Score 21; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.75;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      209 TAGACCTCGCGCGCTGCTACC 229
          |||||
DB      1 TAGACCTCGCGCGCTGCTACC 21
```

```
RESULT 27
US-10-292-198-1
; Sequence 1, Application US/10292198
; Publication No. US20030157654A1
; GENERAL INFORMATION:
; APPLICANT: SHEN, Ben
; APPLICANT: LIU, Wen
; TITLE OF INVENTION: BIOSYNTHEIS OF ENEDIYNE COMPOUNDS BY MANIPULATION OF C-1027 GENE
; FILE REFERENCE: 054030-0007
; CURRENT APPLICATION NUMBER: US/10/292,198
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: US 10/159,257
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 09/478,188
```

```
; PRIOR FILING DATE: 2000-01-05
; PRIOR APPLICATION NUMBER: US 60/115,434
; PRIOR FILING DATE: 1999-01-06
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 63158
; TYPE: DNA
; ORGANISM: Streptomyces globisporus
US-10-292-198-1
```

```
Query Match          3.4%; Score 20; DB 14; Length 63158;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      368 CCGAGGGGCTGCGCGGCTG 387
          |||||
DB      55944 CCGAGGGGCTGCGCGGCTG 55963
```

Search completed: March 12, 2004, 17:42:55
Job time : 195 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 09:38:52 ; Search time 65 seconds
(without alignments)

5097.011 Million cell updates/sec

Title: US-09-966-880A-7_COPY_80_676

Sequence: 1 atgacagacccttcgatgaa.....ttcgacttgggaacttga 597

Scoring table: OLIGO_MTC
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 27747546 residues

Word size: 20

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database:

Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-----------------|
| C 1 | 20 | 3.4 | 1740 | 2 | US-08-796-488-1 |
| C 2 | 20 | 3.4 | 1740 | 3 | US-09-243-934-1 |

ALIGNMENTS

RESULT 1
US-08-796-488-1/c
Sequence 1, Application US/08796488
Patent No. 5928927
GENERAL INFORMATION:
APPLICANT: CHENG, TU-CHEN
APPLICANT: DEFRANK, JOSEPH J.
TITLE OF INVENTION: ENZYMATIC DETOXIFICATION OF
TITLE OF INVENTION: ORGANOPHOSPHORUS COMPOUNDS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: U.S. Army Chemical and Biological Defense
ADDRESS: Command
STREET: Office of the Chief Counsel, Bldg E4435
CITY: Aberdeen Proving Ground
STATE: MD
COUNTRY: U.S.
ZIP: 21010-5423
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/796,488
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bilton, U. J.
REGISTRATION NUMBER: 39,908
REFERENCE/DOCKET NUMBER: DAM 443-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: 410-671-1158
TELEFAX: 410-671-2534
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1740 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Alteromonas sp.
STRAIN: JD6.5
IMMEDIATE SOURCE:
CLONE: pTC6513
FEATURE:
NAME/KEY: -10 signal
LOCATION: 14..18
FEATURE:
NAME/KEY: RBS
LOCATION: 24..30
FEATURE:
NAME/KEY: CDS
LOCATION: 36..1586
US-08-796-488-1

Query Match 3.4%; Score 20; DB 2; Length 1740;
Best Local Similarity 100.0%; Pred. No. 0.84;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 CCTCTGAGACCCCTGCTAC 264 20NT
DB 1605 CCTCTGAGACCCCTGCTAC 1586

RESULT 2
US-09-243-934-1/c
Sequence 1, Application US/09243934
Patent No. 6080566
GENERAL INFORMATION:
APPLICANT: CHENG, TU-CHEN
APPLICANT: DEFRANK, JOSEPH J.
TITLE OF INVENTION: ENZYMATIC DETOXIFICATION OF
TITLE OF INVENTION: ORGANOPHOSPHORUS COMPOUNDS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: U.S. Army Chemical and Biological Defense
ADDRESS: Command
STREET: Office of the Chief Counsel, Bldg E4435
CITY: Aberdeen Proving Ground
STATE: MD
COUNTRY: U.S.
ZIP: 21010-5423
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/243,934

```

/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/796,488
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Biffoni, U. J.
/ REGISTRATION NUMBER: 39,908
/ REFERENCE/DOCKET NUMBER: DAM 443-96
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 410-671-1158
/ TELEFAX: 410-671-2534
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1740 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Alteromonas sp.
/ STRAIN: JDE.5
/ IMMEDIATE SOURCE:
/ CLONE: pTC6513
/ FEATURE:
/ NAME/KEY: -10_signal
/ LOCATION: 14..18
/ FEATURE:
/ NAME/KEY: RBS
/ LOCATION: 24..30
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 36..1586
/ US-09-243-934-1

```

```

Query Match      3.4%; Score 20; DB 3; Length 1740;
Best Local Similarity 100.0%; Pred.No. 0.84;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      245 CCTCCTGAGAGCCCTGCTAC 264
Db      1605 CCTCCTGAGAGCCCTGCTAC 1586

```

Search completed: March 12, 2004, 12:27:53
Job time : 65 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2004, 05:04:59 ; Search time 2369 seconds

(without alignments)
1036.669 Million cell updates/sec

Title: US-09-966-880a-11

Sequence: 1 agagacacatcatatcga.....ctgacacacactatgacag 87

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 20

Total number of hits satisfying chosen parameters: 17

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gse_hum:*
18: em_gse_inv:*
19: em_gse_pin:*
20: em_gse_vrt:*
21: em_gse_fun:*
22: em_gse_mam:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rtd:*
26: em_gse_phg:*
27: em_gse_vr1:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 87 | 100.0 | 535 | 14 | CD707143 |
| 2 | 87 | 100.0 | 693 | 12 | EG757392 |
| 3 | 87 | 100.0 | 1201 | 13 | BK402063 |
| 4 | 84 | 96.6 | 541 | 10 | BF238155 |

| Result | Score | Query Match | Length | ID | Description |
|--------|-------|-------------|--------|----|-------------|
| 5 | 84 | 96.6 | 743 | 12 | BG686133 |
| 6 | 84 | 96.6 | 942 | 10 | BF975166 |
| 7 | 71 | 81.6 | 953 | 13 | B0065440 |
| 8 | 71 | 81.6 | 1052 | 13 | B0055935 |
| 9 | 51 | 58.6 | 820 | 12 | BG757089 |
| 10 | 37 | 42.5 | 889 | 12 | BG686876 |
| 11 | 36 | 41.4 | 872 | 12 | BG758510 |
| 12 | 34 | 39.1 | 1201 | 9 | AL559877 |
| 13 | 26 | 29.9 | 853 | 13 | BK464579 |
| 14 | 21 | 24.1 | 730 | 28 | AQ781292 |
| 15 | 20 | 23.0 | 520 | 9 | A1936315 |
| 16 | 20 | 23.0 | 617 | 13 | BU200032 |
| 17 | 20 | 23.0 | 701 | 13 | BU351369 |

ALIGNMENTS

RESULT 1
LOCUS CD707143 535 bp mRNA linear EST 25-JUN-2003
DEFINITION EST23670 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION CD707143
VERSION CD707143.1 GI:32237773
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
1 (bases 1 to 535)

AUTHORS Zeng,Y.-X.
Transcriptional Gene Expression Profile of Human Nasopharynx

JOURNAL Unpublished (2003)

COMMENT Contact: Yixun Zeng
Cancer Center

Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China

Tel: 86-1380-9770-743
Fax: 86-20-8775-4506

Email: yxzeng@qzsun.edu.cn.
Location/Qualifiers

FEATURES
Source 1..535
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/issue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"

ORIGIN

Query Match 100.0%; Score 87; DB 14; Length 535;
Best Local Similarity 100.0%; Pred. No. 4.5e-37;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| QY | DB | Sequence |
|-----|--|----------|
| 1 | AGAACCATCATTAATTAAGTGAATTTTCTGACCTGAGACTTGACGAGGAGCAGAGA 60 | |
| 42 | AGAACCATCATTAATTAAGTGAATTTTCTGACCTGAGACTTGACGAGGAGCAGAGA 101 | |
| 61 | AGAACCATCATTAATTAAGTGAATTTTCTGACCTGAGACTTGACGAGGAGCAGAGA 87 | |
| 102 | AGAACCATCATTAATTAAGTGAATTTTCTGACCTGAGACTTGACGAGGAGCAGAGA 128 | |

RESULT 2
LOCUS EG757392 693 bp mRNA linear EST 15-MAY-2001
DEFINITION 602711022F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:485180 5',
mRNA sequence.
ACCESSION EG757392
VERSION EG757392.1 GI:14068045
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 693)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LNCM1694 row: K column: 05
High quality sequence stop: 693.
Location/Qualifiers
1..693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4851580"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

ORIGIN
Query Match 100.0%; Score 87; DB 12; Length 693;
Best Local Similarity 100.0%; Pred. No. 4.7e-37;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGACATCATTAATTGAGTGAATTTCTGCGCTGAGACTTGCAGGGAGGCAAGA 60
Db 2 AGAAGACATCATTAATTGAGTGAATTTCTGCGCTGAGACTTGCAGGGAGGCAAGA 61
QY 61 AGAAGACTTGGACACCATATGAGCAG 87
Db 62 AGAAGACTTGGACACCATATGAGCAG 88

RESULT 3
BX402063 1201 bp mRNA linear EST 13-MAY-2003
LOCUS BX402063 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL012YD18 5-PRIME, mRNA sequence.
ACCESSION BX402063
VERSION BX402063.1 GI:30626645
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Tessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr; Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 6672.r For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DL012YD18&cluster=6672.r>. Contact: Feng Liang Email: fliang@life.com URL: <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID: CS0DL012YD18.09PL.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL012YD18"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_1lb="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 100.0%; Score 87; DB 13; Length 1201;
Best Local Similarity 100.0%; Pred. No. 5e-37;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGACATCATTAATTGAGTGAATTTCTGCGCTGAGACTTGCAGGGAGGCAAGA 60
Db 73 AGAAGACATCATTAATTGAGTGAATTTCTGCGCTGAGACTTGCAGGGAGGCAAGA 132
QY 61 AGAAGACTTGGACACCATATGAGCAG 87
Db 133 AGAAGACTTGGACACCATATGAGCAG 159

RESULT 4
BF238155 541 bp mRNA linear EST 14-NOV-2000
LOCUS BF238155 601811880F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4054915 5', mRNA sequence.
DEFINITION BF238155
ACCESSION BF238155
VERSION BF238155.1 GI:11152074
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 541)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LNCM895 row: P column: 20
High quality sequence stop: 541.
Location/Qualifiers
1..541
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4054915"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;

FEATURES
source

Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 96.6%; Score 84; DB 10; Length 541;
 Best Local Similarity 100.0%; Pred. No. 2e-35;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GAACCATCATTAAATGAAGTGAATTTTCTGCGCTGAGACTTGCAGGAGGCAAGAGA 63
 DB 2 GAACCATCATTAAATGAAGTGAATTTTCTGCGCTGAGACTTGCAGGAGGCAAGAGA 61

QY 64 CACTCTGACACCACTATGACAG 87
 DB 62 CACTCTGACACCACTATGACAG 85

RESULT 5 743 bp mRNA linear EST 01-MAY-2001
 BG686133 602638412F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4766234 5',
 LOCUS mRNA sequence.

ACCESSION BG686133
 VERSION BG686133.1 GI:13917530
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgs.nci.nih.gov/.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNCM1626 row: 5 column: 03
 High quality sequence stop: 740.

FEATURES

source

1..743
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4766234"
 /tissue_type="primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_48"
 /note="Organ: B-cells; Vector: pOT7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 96.6%; Score 84; DB 12; Length 743;
 Best Local Similarity 100.0%; Pred. No. 2.1e-35;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GAACCATCATTAAATGAAGTGAATTTTCTGCGCTGAGACTTGCAGGAGGCAAGAGA 63
 DB 2 GAACCATCATTAAATGAAGTGAATTTTCTGCGCTGAGACTTGCAGGAGGCAAGAGA 61

QY 64 CACTCTGACACCACTATGACAG 87
 DB 62 CACTCTGACACCACTATGACAG 85

RESULT 6 942 bp mRNA linear EST 22-JAN-2001
 BP975166 602244657F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:435639 5',
 LOCUS mRNA sequence.

ACCESSION BP975166
 VERSION BP975166.1 GI:12342381
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgs.nci.nih.gov/.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNCM1207 row: a column: 16
 High quality sequence stop: 707.

FEATURES

source

1..942
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:435639"
 /tissue_type="primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_48"
 /note="Organ: B-cells; Vector: pOT7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 96.6%; Score 84; DB 10; Length 942;
 Best Local Similarity 100.0%; Pred. No. 2.2e-35;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GAACCATCATTAAATGAAGTGAATTTTCTGCGCTGAGACTTGCAGGAGGCAAGAGA 63
 DB 2 GAACCATCATTAAATGAAGTGAATTTTCTGCGCTGAGACTTGCAGGAGGCAAGAGA 61

QY 64 CACTCTGACACCACTATGACAG 87
 DB 62 CACTCTGACACCACTATGACAG 85

RESULT 7 953 bp mRNA linear EST 02-APR-2002
 BQ065440 BQ065440

DEFINITION AGENCOURT 6855061 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5929977
5' mRNA sequence.
ACCESSION B065440
VERSION B065440.1 GI:19894486
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 953)
NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LNCM2108 row: P column: 10
High quality sequence stop: 634.
Location/Qualifiers

FEATURES
source
1..953
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5929977"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_99"
/note="Organ: lymph; Vector: pOTB7; site 1: XhoI; site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

ORIGIN

Query Match 81.6%; Score 71; DB 13; Length 953;
Best Local Similarity 100.0%; Pred. No. 3.3e-28;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 TTGAAGTGAATTTTCTGGCTGAGACTTGACGAGGAGGCAAGAGACACTCTGACACC 76
DB 1 TTGAAGTGAATTTTCTGGCTGAGACTTGACGAGGAGGCAAGAGACACTCTGACACC 60

QY 77 ACTATGGACAG 87
DB 61 ACTATGGACAG 71

RESULT 8
LOCUS B0655935 1052 bp mRNA linear EST 29-MAR-2002
DEFINITION AGENCOURT 6796291 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5808181
5' mRNA sequence.
ACCESSION B0655935
VERSION B0655935.1 GI:19815262
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1052)
NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LNCM2051 row: W column: 14
High quality sequence stop: 665.
Location/Qualifiers

FEATURES
source
1..1052
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5808181"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_99"
/note="Organ: lymph; Vector: pOTB7; site 1: XhoI; site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

ORIGIN

Query Match 81.6%; Score 71; DB 13; Length 1052;
Best Local Similarity 100.0%; Pred. No. 3.3e-28;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 TTGAAGTGAATTTTCTGGCTGAGACTTGACGAGGAGGCAAGAGACACTCTGACACC 76
DB 1 TTGAAGTGAATTTTCTGGCTGAGACTTGACGAGGAGGCAAGAGACACTCTGACACC 60

QY 77 ACTATGGACAG 87
DB 61 ACTATGGACAG 71

RESULT 9
LOCUS B6757089 820 bp mRNA linear EST 15-MAY-2001
DEFINITION 602715124F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:485517 5',
mRNA sequence.
ACCESSION B6757089
VERSION B6757089.1 GI:14067742
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 820)
NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Invitrogen Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LNCM1704 row: O column: 06
High quality sequence stop: 675.
Location/Qualifiers

FEATURES
source
1..820

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:485517"
/issue_type="Primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 58.6%; Score 51; DB 12; Length 820;
Best Local Similarity 100.0%; Pred. No. 3.5e-17;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGAGACCATCATTAATGAGTGAATTTCTGCGCTGAGACTTGACAG 51
|||||
6 AGAGACCATCATTAATGAGTGAATTTCTGCGCTGAGACTTGACAG 56

Db 6 AGAGACCATCATTAATGAGTGAATTTCTGCGCTGAGACTTGACAG 56

RESULT 10 889 bp mRNA linear EST 01-MAY-2001
LOCUS BG686876 602650861F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4763247 5'
DEFINITION mRNA sequence.
ACCESSION BG686876
VERSION BG686876.1 GI:13918273
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE NIH-MGC http://mgs.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov

Plate: LNCM1618 row: J column: 16
High quality sequence start: 6
High quality sequence stop: 727.
Location/Qualifiers
1.889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4763247"
/issue_type="Primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Query Match 42.5%; Score 37; DB 12; Length 889;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 51 GGAGGCAAGAGACACTGTGACACCACTATGACAG 87
|||||
Db 9 GGAGGCAAGAGACACTGTGACACCACTATGACAG 45

RESULT 11 872 bp mRNA linear EST 15-MAY-2001
LOCUS BG758510 602712721F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:485069 5'
DEFINITION mRNA sequence.
ACCESSION BG758510
VERSION BG758510.1 GI:14069163
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE NIH-MGC http://mgs.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov

Plate: LNCM1698 row: I column: 06
High quality sequence stop: 836.
Location/Qualifiers
1.872
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4853069"
/issue_type="Primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Query Match 41.4%; Score 36; DB 12; Length 872;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 ATTGAAGTGAATTTCTGCGCTGAGACTTGACAG 51
|||||
Db 2 ATTGAAGTGAATTTCTGCGCTGAGACTTGACAG 37

RESULT 12 1201 bp mRNA linear EST 31-MAY-2003
LOCUS A1559877 1201 bp mRNA linear EST 31-MAY-2003
DEFINITION A1559877 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
clone CSDDG003YB14 5-PRIME, mRNA sequence.
ACCESSION A1559877

VERSION AL559877.2 GI:31284008
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1201)
 AUTHORS Li W.B., Gruber C., Jesse J., and Polayes D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On Feb 15, 2001 this sequence version replaced gi:12905793.
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6672.r For more information about this cluster, see
 http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D6003DA07QPL&cluster=6672.r. Contact : Feng Liang, Email: fliang@life.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0D6003DA07QPL.

FEATURES
 source
 1..1201
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 /clone="CS0D6003YB14"
 /tissue_type="B CELLS (RAMOS CELL LINE)"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN
 Query Match 39.1%; Score 34; DB 9; Length 1201;
 Best Local Similarity 100.0%; Pred. No. 8.9e-08;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 GGCAGAGAGACTGTGACACCACTATGACAG 87
 |||
 66 GGCAGAGAGACTGTGACACCACTATGACAG 99

RESULT 13
 BX464579 853 bp mRNA linear EST 22-MAY-2003
 LOCUS BX464579 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
 DEFINITION clone CS0D6003YB14 5-PRIME, mRNA sequence.
 ACCESSION BX464579
 VERSION BX464579.1 GI:31031641
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 853)
 AUTHORS Li W.B., Gruber C., Jesse J., and Polayes D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6672.r For more information about this cluster, see
 http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS1D6001ZF10QPL&cluster=6672.r. Contact :

Feng Liang Email: fliang@life.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS1D6001ZF10QPL.

FEATURES
 source
 1..853
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 /clone="CS0D6003YB14"
 /tissue_type="B CELLS (RAMOS CELL LINE)"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN
 Query Match 29.9%; Score 26; DB 13; Length 853;
 Best Local Similarity 100.0%; Pred. No. 0.0022;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 GACACTGTGACACCACTATGACAG 87
 |||
 69 GACACTGTGACACCACTATGACAG 94

RESULT 14
 AQ781292 730 bp DNA linear GSS 02-AUG-1999
 LOCUS AQ781292
 DEFINITION HS 3104 B1 B04 MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone plate=3104 Col=7 Row=D, genomic survey sequence.
 ACCESSION AQ781292
 VERSION AQ781292.1 GI:5684252
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 730)
 AUTHORS Mahairas G.G., Wallace J.C., Smith K., Swartzell S., Holzman T., Keller A., Shaker R., Furlong T., Young J., Zhao S., Adams M.D. and Hood L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones may be purchased from Research Genetics (info@resgen.com).
 BAC and Web Server: http://www.htsc.washington.edu
 Plate: 3104 row: D column: 7
 Seq primer: M13 Reverse
 Class: BAC ends
 High quality sequence stop: 730.
 Location/Qualifiers
 1..730
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=3104 Col=7 Row=D"
 /sex="male"
 /note="Organ: sperm; Vector: pBeloBAC11, BAC clones in E-Coli DH10B"

ORIGIN

Query Match 24.1%; Score 21; DB 28; Length 730;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 GGAGCGAAGAGACACTCTGG 71
 |||||
 Db 68 GGAGCGAAGAGACACTCTGG 88

21 X

RESULT 15 A1936315 520 bp mRNA linear EST 08-MAR-2000
 LOCUS A1936315
 DEFINITION w66g09.x1 NCI CGAP Pr22 Homo sapiens cDNA clone IMAGE:2460160 3'
 similar to gb:U63488 REPLICATION PROTEIN A 70 KD DNA-BINDING
 SUBUNIT (HUMAN); mRNA sequence.

ACCESSION A1936315 GI:5675185
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 520)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILM ac:
www-bio.liml.gov/bbrp/image/image.html
 Insert length: 634 Std Error: 0.00
 Seq primer: -40up from Gibco
 High quality sequence stop: 444.

FEATURES

Source

1..520
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2460160"
 /sex="male"
 /tissue_type="normal prostate"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Pr22"
 /note="Organ: prostate; Vector: pRT3D-Pac (Pharmacia)
 with a modified polylinker; 1st strand cDNA was prepared
 from normal prostate bulk tissue, and was then primed with
 a Not I - oligo(dT) primer. Double-stranded cDNA was
 ligated to Eco RI adaptors (Pharmacia), digested with Not
 I and cloned into the Not I and Eco RI sites of the
 modified pRT3 vector. Library is normalized, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 23.0%; Score 20; DB 9; Length 520;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 CATCATTAATTGAGTGAGA 27
 |||||
 Db 159 CATCATTAATTGAGTGAGA 178

RESULT 16
 BU200032/c

LOCUS BU200032 617 bp mRNA linear EST 25-NOV-2002
 DEFINITION 60415206F1 CSECHN03 Gallus gallus cDNA clone CHEST994518 5', mRNA
 sequence.

ACCESSION BU200032
 VERSION BU200032.1 GI:25363376
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 617)
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)

JOURNAL MEDLINE
 PUBMED 12445392
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)

PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES

Source

1..617
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, Hisex"
 /db_xref="taxon:9031"
 /clone="CHEST994918"
 /tissue_type="whole embryo"
 /dev_stage="20-21"
 /lab_host="DH10B"
 /clone_lib="CSECHN03"
 /note="Organ: whole embryo; Vector: pluescript II KS(+);
 Site 1: EcoRI; Site 2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was bluntend, ligated to NotI adaptors, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

ORIGIN

Query Match 23.0%; Score 20; DB 13; Length 617;
 Best Local Similarity 100.0%; Pred. No. 4.4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 TTATTTGAGTGAGATTTT 32
 |||||
 Db 337 TTATTTGAGTGAGATTTT 318

RESULT 17 BU351369 701 bp mRNA linear EST 28-NOV-2002
 LOCUS BU351369
 DEFINITION 603529226F1 CSECHN69 Gallus gallus cDNA clone CHEST48067 5', mRNA
 sequence.

ACCESSION BU351369
 VERSION BU351369.1 GI:25859370
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 701)

Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,

Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.

A Comprehensive Collection of Chicken cDNAs

Curr. Biol. 12 (22), 1965-1969 (2002)

22335534

MEDLINE

PUBMED

COMMENT

Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology

(UMIST)

PO Box 88, Manchester, M60 1QP, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

FEATURES

source

1. 701

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Compton line 15r"

/db_xref="taxon:9031"

/clone="CHEST480d7"

/sex="Female"

/tissue_type="cerebellum"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="CSEQCHN69"

/note="Organ: brain; Vector: pluscript II KS(+); Site_1:

Scor1; Site_2: Not1; This normalized library was

constructed from 1 million independent clones. cDNA

synthesis was initiated using an oligo(dt) primer, using

methylated C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA

was bluntend, ligated to NotI adapters, digested with

EcoRI, size-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the

pluscript (KS+) vector. The library was normalized in 2

rounds using conditions adapted from Soares et al., PNAS

(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6

(1996) 791, except that a significantly longer

reannealing hybridization was used."

ORIGIN

Query Match 23.0%; Score 20; DB 13; Length 701;

Best Local Similarity 100.0%; Pred.No. 4.4;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TTAATTGAAGTGAATTTT 32

DB 128 TTAATTGAAGTGAATTTT 109

Search completed: March 13, 2004, 09:16:27
Job time : 2372 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 12:27:59 ; Search time 944.481 Seconds

(without alignments)
13767.261 Million cell updates/sec

Title: US-09-966-880a-10_COPY_1_300

Perfect score: 300
Sequence: 1 999999cctgtaatccacgct.....atgcgcatggtgacactggt 300

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_bhg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pt:*
10: gb_ro:*
11: gb_sy:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*
15: gb_vt:*
16: em_ba:*
17: em_ba:*
18: em_ba:*
19: em_ba:*
20: em_ba:*
21: em_ba:*
22: em_ba:*
23: em_ba:*
24: em_ba:*
25: em_ba:*
26: em_ba:*
27: em_ba:*
28: em_ba:*
29: em_ba:*
30: em_ba:*
31: em_ba:*
32: em_ba:*
33: em_ba:*
34: em_ba:*
35: em_ba:*
36: em_ba:*
37: em_ba:*
38: em_ba:*
39: em_ba:*
40: em_ba:*
41: em_ba:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|-----------|---------------------|
| 1 | 300 | 100.0 | 6564 | 6 | BD016835 | BD016835 Novel cyt |
| 2 | 298.4 | 99.5 | 11204 | 6 | BD016860 | BD016860 Novel cyt |
| 3 | 298.4 | 99.5 | 11204 | 6 | AB040430 | AB040430 Homo sapi |
| 4 | 298.4 | 99.5 | 71132 | 9 | AC092184 | AC092184 Homo sapi |
| 5 | 109.2 | 36.4 | 79528 | 6 | BD176843 | BD176843 A method |
| 6 | 109.2 | 36.4 | 79528 | 6 | HS466N1 | 297630 Human DNA s |
| 7 | 107.8 | 35.9 | 47177 | 9 | AL157391 | AL157391 Human DNA |
| 8 | 104.6 | 34.9 | 134060 | 9 | AC010389 | AC010389 Homo sapi |
| 9 | 104.6 | 34.9 | 163673 | 9 | AC093240 | AC093240 Homo sapi |
| 10 | 104 | 34.7 | 198521 | 2 | AC145723 | AC145723 Homo sapi |
| 11 | 104 | 34.7 | 201237 | 2 | AC141081 | AC141081 Homo sapi |
| 12 | 104 | 34.7 | 202689 | 2 | AC034281 | AC034281 Homo sapi |
| 13 | 103.6 | 34.5 | 202533 | 2 | AC146464 | AC146464 Salimiri s |
| 14 | 103.2 | 34.4 | 178016 | 2 | AC146467 | AC146467 Actus nan |
| 15 | 102.4 | 34.1 | 161864 | 2 | AC007601 | AC007601 Homo sapi |
| 16 | 102.4 | 34.1 | 174136 | 2 | AC126766 | AC126766 Homo sapi |
| 17 | 101.8 | 33.9 | 161090 | 9 | AC117415 | AC117415 Homo sapi |
| 18 | 101.6 | 33.9 | 131329 | 9 | AC004873 | AC004873 Homo sapi |
| 19 | 101.6 | 33.9 | 157435 | 9 | HS50024 | AL034380 Human DNA |
| 20 | 101.6 | 33.9 | 170926 | 9 | AC095034 | AC095034 Homo sapi |
| 21 | 101.6 | 33.9 | 182703 | 2 | AC022939 | AC022939 Homo sapi |
| 22 | 101.6 | 33.9 | 185045 | 2 | AC146466 | AC146466 Callithr |
| 23 | 101.6 | 33.9 | 185045 | 2 | AC009719 | AC009719 Homo sapi |
| 24 | 101.2 | 33.7 | 153211 | 9 | AC011933 | AC011933 Homo sapi |
| 25 | 101 | 33.7 | 147534 | 9 | AC068620 | AC068620 Homo sapi |
| 26 | 100.8 | 33.6 | 130934 | 9 | EX005428 | EX005428 Human DNA |
| 27 | 100 | 33.3 | 97757 | 9 | AC092402 | AC092402 Homo sapi |
| 28 | 100 | 33.3 | 118047 | 9 | HSB153M9 | AL049631 Human DNA |
| 29 | 99.8 | 33.3 | 61450 | 9 | HS15D7 | AL031229 Human DNA |
| 30 | 99.8 | 33.3 | 109745 | 2 | AC025469 | AC025469 Homo sapi |
| 31 | 99.8 | 33.3 | 121394 | 2 | AL137799 | AL137799 Homo sapi |
| 32 | 99.8 | 33.3 | 171710 | 2 | AC068049 | AC068049 Homo sapi |
| 33 | 99.8 | 33.3 | 204151 | 2 | AC145140 | AC145140 Homo sapi |
| 34 | 99.6 | 33.2 | 192653 | 2 | AC124781 | AC124781 Homo sapi |
| 35 | 99.6 | 33.2 | 209775 | 2 | AC027219 | AC027219 Homo sapi |
| 36 | 99.6 | 33.2 | 246322 | 2 | AC133966 | AC133966 Homo sapi |
| 37 | 99.4 | 33.1 | 129957 | 9 | HS415G2 | Z83846 Human DNA s |
| 38 | 99.2 | 33.1 | 161002 | 2 | AL445664 | AL445664 Human DNA |
| 39 | 99.2 | 33.1 | 217022 | 2 | AC146146 | AC146146 Pan trogl |
| 40 | 99 | 33.0 | 114450 | 9 | AB023058 | AB023058 Homo sapi |
| 41 | 99 | 33.0 | 115133 | 9 | AL669813 | AL669813 Human DNA |
| 42 | 99 | 33.0 | 119030 | 9 | AL844851 | AL844851 Human DNA |
| 43 | 99 | 33.0 | 129818 | 9 | AP000521 | AP000521 Homo sapi |
| 44 | 99 | 33.0 | 166942 | 9 | HS1054A22 | AL031651 Human DNA |
| 45 | 99 | 33.0 | 319486 | 9 | AF055066 | AF055066 Homo sapi |

ALIGNMENTS

RESULT 1
LOCUS BD016835 6564 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel cytidine deaminase.
ACCESSION BD016835
VERSION BD016835.1 GI:22558011
KEYWORDS UP 2001245669-A/8.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 6564)
AUTHORS Honjo T. and Muramatsu M.
TITLE Novel cytidine deaminase
JOURNAL Patent: JP 2001245669-A 8 11-SEP-2001;

| | | |
|----|------------|---|
| | COMMENT | JAPAN TOBACCO INC, TASUKU HONJO OS Homo sapiens (human) PN JP 2001245669-A/8 PD 11-SEP-2001 PF 28-MAR-2000 JP 2000092981 PI TASUKU HONJO, MASAMICHI MORIMATSU PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12, PC A61P7/00, PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC C12N1/21, PC C12N5/10, C12N9/78, C12P21/02, C12P21/08// (C12N1/21, C12R1:19), (C12N5/10, C12R1:91), C12N15/00, C12N5/00, (C12N5/00, C12R1:91) CC |
| | FEATURES | FEATURES SOURCE 1. 6564 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" |
| | ORIGIN | ORIGIN Query Match 100.0%; Score 300; DB 6; Length 6564; Best Local Similarity 100.0%; Pred. No. 3.2e-66; Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| Qy | | 1 GGGGCGCTGTATCCCACTACTCTAGAGGCGTGAAGGCAGAGGAGMTCGCCGAGCCTGSCA 60 |
| Db | | 1 GGGGCGCTGTATCCCACTACTCTAGAGGCGTGAAGGCAGAGGAGATCCCAGAGCCTGSCA 60 |
| Qy | | 61 GATCGCCTTAGAGCCTTGAGAGTTGAGCTAGTAAGCCAAGATCATGCCAGTATCTTC 120 |
| Db | | 61 GATCGCCTTAGAGCCTTGAGAGTTGAGCTAGTAAGCCAAGATCATGCCAGTATCTTC 120 |
| Qy | | 121 AGCCTGGCGCAAAAGTGAGACCGTACAAAAAAAATTAAAAAAGAATTATG 180 |
| Db | | 121 AGCCTGGCGCAAAAGTGAGACCGTACAAAAAAAATTAAAAAAGAATTATG 180 |
| Qy | | 181 ATCAAGATCCCACTGTATAAAAGTGGCCCTMAACAACAATTAAAGTTTGAGTTATTC 240 |
| Db | | 181 ATCAAGATCCCACTGTATAAAAGTGGCCCTMAACAACAATTAAAGTTTGAGTTATTC 240 |
| Qy | | 241 TGCAAGCGAAGAGAACCATCATCGAGGCGCTTTCAGCATGGAGATGGCATGGTGCACCTGGT 300 |
| Db | | 241 TGCAAGCGAAGAGAACCATCATCGAGGCGCTTTCAGCATGGAGATGGCATGGTGCACCTGGT 300 |
| | RESULT 2 | |
| | LOCUS | BD016860 11204 bp DNA linear PAT 27-AUG-2002 |
| | DEFINITION | Novel cytidine deaminase. |
| | ACCESSION | BD016860 |
| | VERSION | BD016860.1 GI:22558036 |
| | KEYWORDS | JP 2001245669-A/33. Homo sapiens (human) |
| | SOURCE | Homo sapiens |
| | ORGANISM | Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. |
| | REFERENCE | 1 (bases 1 to 11204) Honjo, T. and Muramatsu, M. Novel cytidine deaminase. Patent: JP 2001245669-A 33 11-SEP-2001; |
| | AUTHORS | JAPAN TOBACCO INC, TASUKU HONJO |
| | TITLE | OS Homo sapiens (human) |
| | JOURNAL | PN JP 2001245669-A/33 PD 11-SEP-2001 PF 28-MAR-2000 JP 2000092981 PI TASUKU HONJO, MASAMICHI MORIMATSU PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12, PC A61P7/00, PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC C12N1/21, |
| | COMMENT | PC C12N5/10, C12N9/78, C12P21/02, C12P21/08// (C12N1/21, C12R1:19), (C12N5/10, C12R1:91), C12N15/00, C12N5/00, (C12N5/00, C12R1:91) CC FH Location/Qualifiers. |
| | | Key Location/Qualifiers |
| | | Location/Qualifiers |
| | | Location/Qualifiers |

FEATURES

source

1. 11204
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.5%; Score 298.4; DB 6; Length 11204;
Best local Similarity 99.7%; Pred. No. 8.3e-66;
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGGGCCCTGTATATCCACAGTACTACAGAGGCTGAGCAGAGAGATCCCGGAGCCTG3CA 60
DB 5216 GGGGGCCCTGTATATCCACAGTACTACAGAGGCTGAGCAGAGAGATCCCGGAGCCTG3CA 52755
QY 61 GATCTGCCCTGAGCCTGGAGGTTGAGGCTACAGTAAAGCCAAAGATCATGCCAGTACTTC 120
DB 5276 GATCTGCCCTGAGCCTGGAGGTTGAGGCTACAGTAAAGCCAAAGATCATGCCAGTACTTC 53358
QY 121 AGCCTGGCGGCAAGTGAGAGCCCTAACAAAAAATTTAAAAAAGAAATTAG 180
DB 5336 AGCCTGGCGGCAAGTGAGAGCCCTAACAAAAAATTTAAAAAAGAAATTAG 53955
QY 181 ATCAAGATCCACTGTATATAAAGTGCCCTTAAACACCATTAAGAGTTTGAGTTTATTC 240
DB 5396 ATCAAGATCCACTGTATATAAAGTGCCCTTAAACACCATTAAGAGTTTGAGTTTATTC 54555
QY 241 TGCAGGCGAAGAGAACCATCATCGGGGGTCTTCAGCATGGGAATGGCATGGTGCACCTGGT 300
DB 5456 TGCAGGCGAAGAGAACCATCATCGGGGGTCTTCAGCATGGGAATGGCATGGTGCACCTGGT 55155

RESULT 3
AB040430 11204 bp DNA linear PRI 03-OCT-2000

LOCUS
DEFINITION
Homo sapiens A1d gene for activation-induced cytidine deaminase,
complete cds.

AB040430
AB040430.1 GI:9988407
A1d, activation-induced cytidine deaminase.
Homo sapiens (human)

SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
Muto,T., Muramatsu,M., Taniwaki,M., Kinoshita,K. and Honjo,T.
1 (sites)
2 (sites)
3 (bases 1 to 11204)
Muto,T., Muramatsu,M., Taniwaki,M., Kinoshita,K. and Honjo,T.
Direct Submission
Submitted (18-MAR-2000) Tasuku Honjo, Kyoto University, Department
of Medical Chemistry, Faculty of Medicine, Yoshida, Sakyo-ku,
Kyoto, Kyoto 606-8501, Japan (E-mail: honjo@omfour.med.kyoto-u.ac.jp,
Tel:81-75-753-4371 (ex.4371), Fax:81-75-753-4388)
Location/Qualifiers
1. 11204
/organism="Homo sapiens"
/mol_type="genomic DNA"

FEATURES

source

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
20460541
11007475
3 (bases 1 to 11204)
Muto,T., Muramatsu,M., Taniwaki,M., Kinoshita,K. and Honjo,T.
Direct Submission
Submitted (18-MAR-2000) Tasuku Honjo, Kyoto University, Department
of Medical Chemistry, Faculty of Medicine, Yoshida, Sakyo-ku,
Kyoto, Kyoto 606-8501, Japan (E-mail: honjo@omfour.med.kyoto-u.ac.jp,
Tel:81-75-753-4371 (ex.4371), Fax:81-75-753-4388)
Location/Qualifiers
1. 11204
/organism="Homo sapiens"
/mol_type="genomic DNA"

standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
<http://gc.bcm.tmc.edu:8088/quality/info/genbank.annotation.html>.

QUALSTAT-REPORT

FEATURES
source location/Qualifiers

1. 71132
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-438L7"
/complement(1..1448)
/note="overlaps bases 1..1448 of clone AC092490"
/function="clone overlap"
187..408
/standard_name="57233"
439..560
/standard_name="92005"
774..881
/rpt_family="MIR"
903..1190
/rpt_family="Alusq"
1191..1213
/rpt_family="AT_rich"
1591..1807
/standard_name="6612"
1744..1819
/standard_name="8198"
1966..2264
/rpt_family="Alusq"
3296..3328
/rpt_family="(TTTC)n"
4087..4249
/rpt_family="Alusq"
4360..4654
/rpt_family="AluY"
complement(5140..5262)
/rpt_family="FLAM_C"
5275..5304
/rpt_family="AT_rich"
5305..5563
/rpt_family="AluDo"
5564..5599
/rpt_family="AT_rich"
5644..5733
/rpt_family="MSTD"
5734..6012
/rpt_family="Alusq"
6020..6336
/rpt_family="Alusq"
6337..6629
/rpt_family="MSTD"
6630..6650
/rpt_family="(CAA)n"
complement(6711..6817)
/rpt_family="MIR"
complement(7033..7327)
/rpt_family="Alusq"
7607..7745
/rpt_family="FLAM_A"
7980..8047
/rpt_family="MIR"
complement(8050..8549)

/rpt_family="LTR47A"
8567..8698
/rpt_family="MIR"
complement(8815..9101)
/rpt_family="Alusq"
complement(11227..12175)
/rpt_family="LTR5"
12625..12900
/standard_name="87432"
complement(13547..13664)
/rpt_family="LIMB5"
complement(13665..13979)
/rpt_family="AluY"
complement(13980..14193)
/rpt_family="LIMB5"
14622..14924
/rpt_family="AluY"
complement(15813..17136)
/rpt_family="LIPAB"
complement(17272..17395)
/rpt_family="FLAM_C"
17426..17575
/rpt_family="MIR"

Query Match 99.5%; Score 298.4; DB 9; Length 71132;
Best Local Similarity 99.7%; Pred. No. 9e-66; 1; Indels 0; Gaps 0;
Matches 299; Conservative 0; Mismatches 1;

QY 1 GGGGGCTGTATTCAGCTACTCAGAGGCTGAGGAGATCCGGAGCTGGCA 60
DB 39760 GGGGGCTGTATTCAGCTACTCAGAGGCTGAGGAGATCCGGAGCTGGCA 39819
QY 61 GATTCGCTGAGCTGGAGGCTTACGCTACAGTAAGCCAAATGCTGCAATCTTC 120
DB 39820 GATTCGCTGAGCTGGAGGCTTACGCTACAGTAAGCCAAATGCTGCAATCTTC 39879
QY 121 AGCTGGGCGCAAGAGTGAACGCTAACCAAAAAAATTTAAAGAAATTTAG 180
DB 39880 AGCTGGGCGCAAGAGTGAACGCTAACCAAAAAAATTTAAAGAAATTTAG 39939
QY 181 ATCAAGATCCAACTGTAAAAAGTGGCTTAAACCAACATTAAAGTTGAGTTATTC 240
DB 39940 ATCAAGATCCAACTGTAAAAAGTGGCTTAAACCAACATTAAAGTTGAGTTATTC 39999
QY 241 TGCAGGCAAGAAAGAACATCAGGGGGCTTTCAGCATGGGAATGGCATGCTGCT 300
DB 40000 TGCAGGCAAGAAAGAACATCAGGGGGCTTTCAGCATGGGAATGGCATGCTGCT 40059

RESULT 5
BD176843/c 79528 bp DNA linear PAT 18-MAR-2003
LOCUS BD176843
DEFINITION A method of predicting cancer condition.
ACCESSION BD176843.1 GI:29122555
VERSION WO 02072828-A/6
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 79528)
Kato, K., Iwao, K., Noguchi, S. and Matsuda, R.
A method of predicting cancer condition
Patent: WO 02072828-A 6 19-SEP-2002;
DNA CHIP RESEARCH INC. HITACHI SOFTWARE ENGINEERING CO LTD, KIKUYA
KATO, KYOKO IMAO, SHINZABURO NOGUCHI, RIO MATOBA
OS Homo sapiens (human)
PN WO 02072828-A/6
PD 19-SEP-2002
PF 07-MAR-2002 WO 2002JP002153
PR 14-MAR-2001 JP 01P 073063, 06-APR-2001 JP 01P 108503 PR
02-AUG-2001 JP 01P 234807
PI KIKUYA KATO, KYOKO IMAO, SHINZABURO NOGUCHI, RIO MATOBA PC

C12N15/12, C12Q1/68, G06F19/00
CC A method of predicting cancer condition
FH Key Location/Qualifiers
FT source 1..79528 /organism='Homo sapiens (human)'.
source Location/Qualifiers
1..79528 /organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Query Match 36.4%; Score 109.2; DB 6; Length 79528;
Best Local Similarity 72.7%; Pred. No. 1,9e-17;
Matches 141; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1 GGGGGGCTGTATTCACAGCTACTCAGAGAGGCTGAGGAGAGATCCGGAGCCTGGCA 60
64094 GGGGGGCTGTATTCACAGCTACTCAGAGAGGCTGAGGAGAGATTCGTAAGCA 64035
QY 61 GATCTGCTTACAGCTGAGGAGGCTGAGGCTACAGTACCCAAATCATGCCAGTACTTC 120
64034 GAATTCGCTTACAGCTGAGGAGGCTGAGGCTACAGTACCCAAATCATGCCAGTACTTC 63975
QY 121 AGGCTGGGAGCAAGTGAAGACGCTAACAAAAAATTAAAAAAGAAATTAG 180
63974 AGGCTGGGAGCAAGTGAAGACGCTAACAAAAAATTAAAAAAGAAATTAG 63915
QY 181 ATCAGATCCACT 194
Db 63914 ATAAATTAATTAAT 63901

RESULT 6
HS466N1/c
LOCUS
DEFINITION
79528 bp. DNA linear PRI 05-JUN-2003
Human DNA sequence from clone RP3-466N1 on chromosome 22q12-13
Contains the H1FO gene for H1 histone family member 0, the GCAT
gene for glycine C-acetyltransferase (2-amino-3-ketobutyrate
coenzyme A ligase), the GALR3 gene for galactin receptor, the gene
for a novel protein similar to ANK3 (ankyrin 3, node of Ranvier
(ankyrin G)), the 5' end of the gene for proteins HSPC025 and
HSPC021 (similar to C. elegans FAT-3 alcohol dehydrogenase), ESTs,
STSs, GSSs and eight putative CpG islands, complete sequence.
297630
297630.11 GI:4582128
HNG; alcohol dehydrogenase; ANK3; ankyrin; CpG island; FAT-3;
galactin receptor; GALR3; GCAT; H1FO; histone; HSPC021; HSPC025.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 79528)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES

source

Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
This sequence was generated from part of bacterial clone contigs of
human chromosome 22, constructed by the Sanger Centre Chromosome 22
Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr22>
RP3-466N1 is from the library RP3-3 constructed by the group of
Pleier de Jong. For further details see
<http://www.choi.org/bacpac/home.htm>
VECTOR: pCYPAC2
IMPORTANT: This sequence is not the entire insert of clone
RP3-466N1. It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true left end of clone RP5-101413 is at 79429 in this
sequence. The true right end of clone RP1-37816 is at 100 in this
sequence.

repeat_region Location/Qualifiers
1..79528 /organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
/chromosome='22'
/map='q12-13'
/clone='RP3-466N1'
/clone_id='RP3-3'
8..17
/note='2.0 copies 5 mer AGCTC 20% conserved'
39..174
/note='AluSg/x repeat: matches 1..136 of consensus'
175..492
/note='AluSx repeat: matches 1..312 of consensus'
494..505
/note='3.0 copies 4 mer GAA 24% conserved'
850..1133
/note='AluJo repeat: matches 22..312 of consensus'
1134..1446
/note='AluSq repeat: matches 1..313 of consensus'
1514..1597
/note='WIR repeat: matches 180..262 of consensus'
1837..2110
/note='AluJo repeat: matches 13..294 of consensus'
2113..2253
/note='F1AM C repeat: matches 1..133 of consensus'
2261..2280
/note='3.3 copies 6 mer AAACAC 26% conserved'
2299..2319
/note='2.3 copies 9 mer TTGATCAG 33% conserved'
2537..2555
/note='3.8 copies 5 mer TTGT 38% conserved'
/note='AluSx repeat: matches 146..311 of consensus'
/note='AluYc2 repeat: matches 1..310 of consensus'
/note='AluYc2 repeat: matches 1..310 of consensus'
/note='AluSx repeat: matches 4..146 of consensus'
3205..3215
/note='11.0 copies 1 mer T 22% conserved'
/note='AluSx repeat: matches 1..312 of consensus'
3555..3766
/note='12 repeat: matches 2443..2691 of consensus'
3842..3908
/note='L1M2 repeat: matches 5672..5727 of consensus'
3909..4040
repeat_region

/note="FLAM_C repeat: matches 3. .133 of consensus"
4041. .4072
repeat_region
/note="LIME repeat: matches 5727. .5761 of consensus"
complement(4073. .4399)
repeat_region
/note="MLTID repeat: matches 162. .505 of consensus"
4400. .4697
repeat_region
/note="AluSp repeat: matches 1. .299 of consensus"
complement(4698. .4860)
repeat_region
/note="MLTID repeat: matches 4. .162 of consensus"
4861. .4886
repeat_region
/note="2.2 copies 12 mer CACACAGTACC 52% conserved"
4897. .4911
repeat_region
/note="2.5 copies 6 mer GGGCTG 30% conserved"
5073. .5113
repeat_region
/note="L1MA1 repeat: matches 6214. .6256 of consensus"
5114. .5420
repeat_region
/note="AluSx repeat: matches 1. .307 of consensus"
5421. .5473
repeat_region
/note="L1MA1 repeat: matches 6256. .6302 of consensus"
5476. .5498
repeat_region
/note="2.3 copies 1 mer A 28% conserved"
complement(5503. .5977)
repeat_region
/note="L2 repeat: matches 3136. .3220 of consensus"
6050. .6393
repeat_region
/note="MER83 repeat: matches 1. .355 of consensus"
complement(6394. .6705)
repeat_region
/note="AluY repeat: matches 1. .311 of consensus"
complement(6706. .7009)
repeat_region
/note="AluSx repeat: matches 1. .305 of consensus"
7010. .7126
repeat_region
/note="MER83 repeat: matches 353. .448 of consensus"
complement(7145. .7445)
repeat_region
/note="L1MA4 repeat: matches 5936. .6236 of consensus"
7446. .7455
repeat_region
/note="10.0 copies 1 mer T 20% conserved"
complement(7466. .7588)
repeat_region
/note="AluY repeat: matches 1. .123 of consensus"
complement(7589. .7887)
repeat_region
/note="AluSp repeat: matches 1. .302 of consensus"
complement(7898. .8117)
repeat_region
/note="AluSg/X repeat: matches 83. .300 of consensus"
complement(8121. .8453)
repeat_region
/note="L1MA4 repeat: matches 5618. .5949 of consensus"
8456. .8510
repeat_region
/note="AluSg/X repeat: matches 119. .133 of consensus"
8511. .8807
repeat_region
/note="AluSx repeat: matches 1. .298 of consensus"
8808. .8983
repeat_region
/note="AluSg/X repeat: matches 133. .306 of consensus"
8995. .9004
repeat_region
/note="3.3 copies 3 mer ACC 20% conserved"
9031. .9082
repeat_region
/note="L2 repeat: matches 3106. .3160 of consensus"
complement(9083. .9386)
repeat_region
/note="AluSq repeat: matches 1. .304 of consensus"
9387. .9522
repeat_region
/note="L2 repeat: matches 3160. .3282 of consensus"
9546. .9852
repeat_region
/note="AluSg1 repeat: matches 1. .306 of consensus"
9947. .9960
repeat_region
/note="2.3 copies 6 mer GGGGTG 28% conserved"
complement(10080. .10164)
repeat_region
/note="MIR repeat: matches 109. .206 of consensus"
10110. .10367
misc_feature
/note="match: GSS: Em:AQ416488"
10157. .10649
misc_feature
/note="CpG island"
/evidence=not_experimental
repeat_region
10299. .10316
repeat_region
/note="2.2 copies 8 mer CCGCCGTT 27% conserved"
10478. .10512
repeat_region
/note="7.0 copies 5 mer GAGGC 61% conserved"
10777. .10788

Query Match 36.4%; Score 109.2; DB 9; Length 79528;
Best Local Similarity 72.7%; Pred. No. 1.9e-17;
Matches 141; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1 GGGGGCCCTGATCCAGCTACTGAGAGCGTGAAGAGAGATCCGGAGCGTGGCA 60
DB 64094 GGGGGCCCTGATCCAGCTACTGAGAGCGTGAAGAGAGATCCGTAACCTGGA 64035
QY 61 GATCGCTGAGCTGGAGGTTGAGGCTACAGTAAGCAAGATCATGCGATATCTTC 120
DB 64034 GATGCTTGAACCTGGAGGCGAGGTTGAGTGAACCAAGTTGCGCACATGATTC 63975
QY 121 AGCCTGGCGCAAGAGTGAAGCCGTAACAAAAAATTTAAAGAAATTTAG 180
DB 63974 AGCCTGGAGAGACAGAGACTCATCTCAAAAAATATATATATATATATATTA 63915
QY 181 ATCAAGATCCACT 194
DB 63914 ATTAATTAATTAAT 63901

RESULT 7
AL157391/c
LOCUS Human DNA sequence from clone RP11-271M1 on chromosome 10, complete
DEFINITION
ACCESSION AL157391
VERSION AL157391.11 GI:15149560
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 47177)
REFERENCE
AUTHORS Chapman,J.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT
requests: clonerequests@sanger.ac.uk
On Aug 13, 2001 this sequence version replaced gi:15021290.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMBASE; information on the WORMBASE

database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr10>
RP11-271M1 is from the library RP11-11.1 constructed by the group of Peter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-271M1 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
The true left end of clone RP11-271M1 is at 1 in this sequence. The true left end of clone RP11-455B2 is at 45178 in this sequence.

FEATURES

Source
1..47177
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone_id="RP11-271M1"
/clone_id="RP11-11.1"
1..2296
/note="HERV repeat: matches 1108..3423 of consensus"
2313..2628
repeat_region
/note="4 copies 79 mer 67% conserved"
2726..2883
repeat_region
/note="2 copies 79 mer 82% conserved"
2797..3100
repeat_region
/note="4 copies 76 mer 70% conserved"
3348..3581
repeat_region
/note="3 copies 78 mer 79% conserved"
3371..3895
misc_feature
/note="CpG island"
/evidence="not experimental"
3499..3806
repeat_region
/note="4 copies 77 mer 75% conserved"
3616..3843
repeat_region
/note="3 copies 76 mer 78% conserved"
3734..3889
repeat_region
/note="2 copies 78 mer 82% conserved"
3949..4134
repeat_region
/note="3 copies 62 mer 75% conserved"
4101..4352
repeat_region
/note="4 copies 63 mer 77% conserved"
4290..4413
repeat_region
/note="2 copies 62 mer 83% conserved"
4525..4902
repeat_region
/note="6 copies 63 mer 77% conserved"
4813..5122
repeat_region
/note="5 copies 62 mer 68% conserved"
4947..5072
repeat_region
/note="2 copies 63 mer 85% conserved"
5164..5415
repeat_region
/note="4 copies 63 mer 75% conserved"
5501..5815
repeat_region
/note="5 copies 63 mer 75% conserved"
5772..6005
repeat_region
/note="3 copies 78 mer 85% conserved"
6094..6245
repeat_region
/note="2 copies 76 mer 85% conserved"
6294..6445
repeat_region
/note="2 copies 76 mer 86% conserved"
6294..6445
misc_feature
/note="match: STS: Em:HSFH09B8"
8157..8457
repeat_region
/note="AluSg repeat: matches 1..299 of consensus"
8524..8636
repeat_region
/note="L1 repeat: matches 2596..2710 of consensus"
9039..9348
repeat_region
/note="AluSg repeat: matches 1..308 of consensus"
9743..9955

repeat_region
/note="MIR repeat: matches 32..252 of consensus"
10799..11045
repeat_region
/note="L1MB3 repeat: matches 5896..6151 of consensus"
11046..11335
repeat_region
/note="AluSx repeat: matches 12..304 of consensus"
11336..11461
repeat_region
/note="L1MB3 repeat: matches 5766..5896 of consensus"
11462..11771
repeat_region
/note="AluSx repeat: matches 6..312 of consensus"
11772..12766
repeat_region
/note="L1MB3 repeat: matches 4818..5766 of consensus"
12812..13119
repeat_region
/note="AluSg repeat: matches 1..308 of consensus"
13200..13406
repeat_region
/note="L1MA4 repeat: matches 5961..6130 of consensus"
13407..13686
repeat_region
/note="AluSx repeat: matches 1..280 of consensus"
13687..13865
repeat_region
/note="L1MA4 repeat: matches 5771..5961 of consensus"
14056..14388
repeat_region
/note="AluSg repeat: matches 1..312 of consensus"
14547..14692
repeat_region
/note="MER5B repeat: matches 25..172 of consensus"
14798..14995
repeat_region
/note="MER6B repeat: matches 5..210 of consensus"
complement(15721..16199)
/note="match: STS: Em:G55345"
16895..17117
repeat_region
/note="AluSg repeat: matches 1..222 of consensus"
17114..17365
repeat_region
/note="4 copies 63 mer 86% conserved"
17293..17444
repeat_region
/note="2 copies 76 mer 83% conserved"
17389..17514
repeat_region
/note="2 copies 63 mer 83% conserved"
17446..17597
repeat_region
/note="2 copies 76 mer 87% conserved"
17528..17716
repeat_region
/note="3 copies 63 mer 89% conserved"
17645..17796
repeat_region
/note="2 copies 76 mer 89% conserved"
17721..17972
repeat_region
/note="4 copies 63 mer 88% conserved"
17913..18444
repeat_region
/note="7 copies 76 mer 82% conserved"
19048..19883
repeat_region
/note="11 copies 76 mer 63% conserved"
19355..19654
repeat_region
/note="2 copies 150 mer 86% conserved"
19645..19800
repeat_region
/note="2 copies 78 mer 82% conserved"
20048..20351
repeat_region
/note="4 copies 76 mer 79% conserved"
20414..20641
repeat_region
/note="3 copies 76 mer 71% conserved"
20697..22765
repeat_region
/note="HERV repeat: matches 3122..5178 of consensus"
23280..23589
repeat_region
/note="AluSx repeat: matches 1..309 of consensus"
23984..24293
repeat_region
/note="AluSg repeat: matches 2..311 of consensus"
24761..25049
repeat_region
/note="AluSx repeat: matches 4..296 of consensus"
25132..25437
repeat_region
/note="AluSx repeat: matches 1..307 of consensus"
25926..26093
repeat_region
/note="AluSg/x repeat: matches 143..309 of consensus"
26150..26300
repeat_region
/note="AluY repeat: matches 135..302 of consensus"
26303..26691
repeat_region
/note="L1MB6 repeat: matches 5700..6095 of consensus"
26722..27017
repeat_region
/note="AluSx repeat: matches 1..298 of consensus"

| | |
|-----------|--|
| REFERENCE | 1 (bases 1 to 163673) |
| AUTHORS | DOE Joint Genome Institute and Stanford Human Genome Center. |
| TITLE | Direct Submmission |
| JOURNAL | Unpublished |
| REFERENCE | 2 (bases 1 to 163673) |
| AUTHORS | DOE Joint Genome Institute. |
| TITLE | Direct Submmission |
| JOURNAL | Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint |
| REFERENCE | 3 (bases 1 to 163673) |
| AUTHORS | Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA |
| TITLE | DOE Joint Genome Institute and Stanford Human Genome Center. |
| JOURNAL | Direct Submmission |
| COMMENT | Submitted (31-JAN-2002) DOE Joint Genome Institute, 2800 Mitchell |
| | Drive, Walnut Creek, CA 94598, USA |
| | On Jan 31, 2002 this sequence version replaced gi:15193374. |
| | Draft Sequence Produced by DOE Joint Genome Institute |
| | www.jgi.doe.gov |
| | Finishing Completed at Stanford Human Genome Center |
| | www.sngc.stanford.edu |
| | Quality: Phrap Quality >=40 99.9% of Sequence; |
| | Estimated Total Number of Errors is 0.2. |
| FEATURES | Location/Qualifiers |
| source | 1..163673 |
| | /organism="Homo sapiens" |
| | /mol_type="genomic DNA" |
| | /db_xref="taxon:9606" |
| | /chromosome="5" |
| | /clone="RP11-103P4" |


```
* 76134 79417: contig of 3284 bp in length
* 79418 79517: gap of unknown length
* 79518 81816: contig of 2299 bp in length
* 81817 81916: gap of unknown length
* 81917 84825: contig of 2909 bp in length
* 84826 84925: gap of unknown length
* 84926 87053: contig of 2128 bp in length
* 87054 87153: gap of unknown length
* 87154 90407: contig of 3254 bp in length
* 90408 90507: gap of unknown length
* 90508 94550: contig of 4043 bp in length
* 94551 94651: gap of unknown length
* 94651 96993: contig of 2343 bp in length
* 96994 97094: gap of unknown length
* 101224 101323: contig of 4130 bp in length
* 101324 104392: contig of 3069 bp in length
* 104393 104492: gap of unknown length
* 104493 108893: contig of 4401 bp in length
* 108894 108993: gap of unknown length
* 108994 113350: contig of 4357 bp in length
* 113351 113450: gap of unknown length
* 113451 118703: contig of 5253 bp in length
* 118704 118803: gap of unknown length
* 118804 121418: contig of 2615 bp in length
* 121419 121518: gap of unknown length
* 121519 126615: contig of 5097 bp in length
* 126616 126715: gap of unknown length
* 126716 131029: contig of 4314 bp in length
* 131030 131129: gap of unknown length
* 131130 137699: contig of 6570 bp in length
* 137700 137799: gap of unknown length
* 137800 143730: contig of 5931 bp in length
* 143731 143830: gap of unknown length
* 143831 151122: contig of 7292 bp in length
* 151123 151223: gap of unknown length
* 151224 159620: contig of 8398 bp in length
* 159621 159720: gap of unknown length
* 159721 172370: contig of 12550 bp in length
* 172371 198521: gap of unknown length
* 172371 198521: contig of 26151 bp in length.
```

FEATURES
source
1..198521 location/Qualifiers

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="XXfos-8909F12"

Query Match 34.7%; Score 104; DB 2; Length 198521;
Best Local Similarity 72.8%; Pred. No. 4.3e-16;
Matches 134; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

```
QY 1 GGGGGCTGTATCCAGCTACTCAGAGGCTGAGGAGGATCCGCGAGCTCGCA 60
DB 13253 GGGTGGCTGTATCCAGCTACTCAGAGGCTGAGGAGGATCCGCTTGAGGAGGA 13194
QY 61 GATTCGCTGAGCTGGAGGTTGAGGCTACAGTAACCAAGTCAATGCCAGTATCTTC 120
DB 13193 GAATCGCTTGAACCCAGAGGCGGAGGTTGCAATGACCCAGAGTCCGCTGCACTCC 13134
QY 121 AGCTGGGCGACAAGTGAAGCCGTAAACAAAAAATTTAAAAAAGAAATTTAG 180
DB 13133 AGCTGGGCGACAAGTGAAGCTTATCTCAAAATAAACAAGTTAATAAATAAAAA 13074
QY 181 ATCA 184
DB 13073 ATTA 13070
```

RESULT 11
AC141081

LOCUS AC141081 201237 bp DNA linear HTG 07-MAR-2003
DEFINITION Homo sapiens chromosome 16 clone RP11-927N1, WORKING DRAFT
SEQUENCE, 3 unordered pieces.
AC141081
ACCESSION AC141081.1 GI:26875941
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 201237)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished
2 (bases 1 to 201237)
DOE Joint Genome Institute.
Direct Submission
Submitted (07-MAR-2003) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 16S0785
Center clone name: RP11-927N1

Summary Statistics
Consensus quality: 200303 bases at least Q40
Consensus quality: 200551 bases at least Q30
Consensus quality: 200605 bases at least Q20
Estimated insert size: 175000; agarose-fp estimation
Estimated insert size: 201037; sum-of-contigs estimation
Quality coverage: 12.74 in Q20 bases; agarose-fp estimation
Quality coverage: 11.09 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 34897: contig of 34897 bp in length
* 34998 34997: gap of unknown length
* 34998 88292: contig of 53295 bp in length
* 88293 88392: gap of unknown length
* 88393 201237: contig of 112845 bp in length.

FEATURES
source

1..201237 location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-927N1"
/clone_id="RP11 human BAC library 11"

ORIGIN

Query Match 34.7%; Score 104; DB 2; Length 201237;
Best Local Similarity 72.8%; Pred. No. 4.3e-16;
Matches 134; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

```
QY 1 GGGGGCTGTATCCAGCTACTCAGAGGCTGAGGAGGATCCGCGAGCTCGCA 60
DB 144542 GGGTGGCTGTATCCAGCTACTCAGAGGCTGAGGAGGATCCGCTTGAGGAGGA 144601
QY 61 GATTCGCTGAGCTGGAGGTTGAGGCTACAGTAACCAAGTCAATGCCAGTATCTTC 120
DB 144602 GAATCGCTTGAACCCAGAGGCGGAGGTTGCAATGACCCAGAGTCCGCTGCACTCC 144661
QY 121 AGCTGGGCGACAAGTGAAGCCGTAAACAAAAAATTTAAAAAAGAAATTTAG 180
```

Db 144662 AGCCTGGGCAACAGAGTGAAGTCTATCTCAATAAACAAGATTAAAGTAATAAAA 144721
 AC034281
 LOCUS 181 ATCA 184
 QY 144722 ATTA 144725
 Db 144722 ATTA 144725

RESULT 12
 AC034281
 LOCUS 181 ATCA 184
 QY 144722 ATTA 144725

DEFINITION Homo sapiens chromosome 16 clone RP11-65606 map 16, WORKING DRAFT

ACCESSION AC034281
 VERSION AC034281.3 GI:10047874
 KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 202689)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS Birren, B., Linton, J., Nusbaum, C. and Lander, E.
 TITLE Homo sapiens chromosome 16, clone RP11-65606
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 202689)
 AUTHORS Birren, B., Linton, J., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F., Boguslavskiy, L., Boukhalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Larocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Locke, K., MacDonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPherson, R., Meldrum, J., Menus, L., Mihova, T., Miranda, C., Mienna, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., O'Neil, T.M., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.U., Young, G., Zainum, J., Zimmer, A. and Zody, M.
 TITLE Direct Submission
 JOURNAL Submitted (05-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 202689)
 AUTHORS Birren, B., Linton, J., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Barna, N., Bastien, V., Beda, F., Boguslavskiy, L., Boukhalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Larocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacDonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McPherson, R., Meldrum, J., Menus, L., Mihova, T., Miranda, C., Mienna, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., O'Neil, T.M., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.U., Young, G., Zainum, J., Zimmer, A. and Zody, M.
 TITLE Direct Submission
 JOURNAL Submitted (24-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Sep 10, 2000 this sequence version replaced gi:7770529.

All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu

Project Information
 Center project name: 18279
 Center clone name: 6560_6

Sequencing Statistics
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 182566 bases at least Q40
 Consensus quality: 192896 bases at least Q30
 Consensus quality: 196996 bases at least Q20
 Insert size: 210000; agarose-fp
 Insert size: 199989; sum-of-contigs
 Quality coverage: 3.9 in Q20 bases; agarose-fp
 Quality coverage: 4.1 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 28 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

| | | |
|--------|---------|------------------------------|
| 1 | 216: | contig of 216 bp in length |
| 217 | 316: | gap of 100 bp |
| 317 | 1822: | contig of 1506 bp in length |
| 1823 | 1922: | gap of 100 bp |
| 1923 | 2968: | contig of 1046 bp in length |
| 2969 | 3068: | gap of 100 bp |
| 3069 | 4380: | contig of 1312 bp in length |
| 4381 | 4480: | gap of 100 bp |
| 4481 | 6135: | contig of 1655 bp in length |
| 6136 | 6235: | gap of 100 bp |
| 6236 | 9213: | contig of 2978 bp in length |
| 9214 | 9313: | gap of 100 bp |
| 9314 | 11857: | contig of 2544 bp in length |
| 11858 | 11957: | gap of 100 bp |
| 11958 | 14517: | contig of 2560 bp in length |
| 14518 | 14617: | gap of 100 bp |
| 14618 | 32144: | contig of 17527 bp in length |
| 32145 | 32244: | gap of 100 bp |
| 32245 | 35856: | contig of 3412 bp in length |
| 35857 | 35756: | gap of 100 bp |
| 35757 | 40409: | contig of 4653 bp in length |
| 40410 | 40509: | gap of 100 bp |
| 40510 | 46096: | contig of 5587 bp in length |
| 46097 | 46196: | gap of 100 bp |
| 46197 | 50343: | contig of 4147 bp in length |
| 50344 | 50443: | gap of 100 bp |
| 50444 | 54901: | contig of 4458 bp in length |
| 54902 | 55001: | gap of 100 bp |
| 55002 | 59436: | contig of 4435 bp in length |
| 59437 | 59536: | gap of 100 bp |
| 59537 | 66852: | contig of 7316 bp in length |
| 66853 | 66952: | gap of 100 bp |
| 66953 | 75809: | contig of 8857 bp in length |
| 75810 | 75909: | gap of 100 bp |
| 75910 | 85009: | contig of 9100 bp in length |
| 85010 | 85109: | gap of 100 bp |
| 85110 | 94166: | contig of 9657 bp in length |
| 94167 | 94266: | gap of 100 bp |
| 94267 | 104445: | contig of 10179 bp in length |
| 104446 | 104545: | gap of 100 bp |
| 104546 | 114898: | contig of 10353 bp in length |
| 114899 | 114998: | gap of 100 bp |

114999 126032: contig of 11034 bp in length
126033 126132: gap of 100 bp
126133 139665: contig of 13533 bp in length
139665 139765: gap of 100 bp
139765 151421: contig of 11656 bp in length
151421 151521: gap of 100 bp
151521 169514: contig of 17993 bp in length
169514 169515: gap of 100 bp
169515 183307: contig of 13693 bp in length
183307 183407: gap of 100 bp
183407 200849: contig of 17442 bp in length
200849 200850: gap of 100 bp
200850 202689: contig of 1740 bp in length.

FEATURES
SOURCE

1. 202689
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16"
/clone="RP11-656C6"
/clone_id="RP11 Human Male BAC"
1. .216
/note="assembly_fragment
clone end:SP6
vector side:left"
misc_feature
317..1822
/note="assembly_fragment"
1923..2968
/note="assembly_fragment"
3069..4380
/note="assembly_fragment"
4481..6135
/note="assembly_fragment"
6236..9213
/note="assembly_fragment"
9314..11857
/note="assembly_fragment"
11958..14517
/note="assembly_fragment"
14618..32144
/note="assembly_fragment"
32245..35656
/note="assembly_fragment"
35757..40409
/note="assembly_fragment"
40510..46096
/note="assembly_fragment"
46197..50343
/note="assembly_fragment"
50444..54901
/note="assembly_fragment"
55002..59436
/note="assembly_fragment"
59537..66852
/note="assembly_fragment"
66953..75809
/note="assembly_fragment"
75910..85009
/note="assembly_fragment"
85110..94166
/note="assembly_fragment"
94267..104445

Query Match 34.7%; Score 104; DB 2; Length 202689;

Best Local Similarity 72.8%; Pred. No. 4,36-16; Indels 0; Gaps 0;

Matches 134; Conservative 0; Mismatches 50;

QY 1 GGGGGCCGTATCCAGCTACTCAGAGGCTGAGGAGGATCCGCGAGGCTTGCA 60
DB 14920 GGGGGCCGTATCCAGCTACTCAGAGGCTGAGGAGGATCCGCGAGGCTTGCA 14979
QY 61 GATCTGCTGAGCTGGAGGCTGAGGCTACAGTAAAGCAAGATCATGCTGCTGCTTCC 120

DB 14980 GAATCGCTTGAACCCAGAGGCGAGGTTGCAATGACCGAGATCGCACTGCACTCC 15039
QY 121 AGCCCTGGGCGCAAGAGTGAGACCGTAAACAAAAAATTTAAAGAAATTTAG 180
DB 15040 AGCCCTGGGCGCAAGAGTGAGACCGTATCTCAAAATTAACAGATTAAAGTAAATAAA 15099
QY 181 ATCA 184
DB 15100 ATTA 15103

RESULT 13

AC146464

LOCUS 202533 bp DNA linear HTG 19-AUG-2003

DEFINITION Saimiri sciureus clone CH254-84A11, WORKING DRAFT SEQUENCE.

AC146464

VERSION AC146464.1 GI:33636782

KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.

SOURCE Saimiri sciureus (common squirrel monkey)

ORGANISM Saimiri sciureus

Bukarjota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae;

Saimiri.

1 (bases 1 to 202533)

Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,

Peng, Z., Malinov, I. and Rubin, E.M.

Direct Submission

Submitted (14-AUG-2003) Genome Sciences, Lawrence Berkeley National

Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA

3 (bases 1 to 202533)

Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,

Peng, Z., Malinov, I. and Rubin, E.M.

Direct Submission

Submitted (19-AUG-2003) Genome Sciences, Lawrence Berkeley National

Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA

Bac Clone Name: CH254-84A11

Sequence Produced by Berkeley PGA

Web site: <http://pga.lbl.gov>

Center Code: PGASERK

Center Project Name: S030

This sequence has been compared to sequences of other species

using VISTA (<http://www.gscl.lbl.gov/VISTA/>). The results can be

viewed at:

http://pga.lbl.gov/cgi-bin/search_cvcgdtvtype=navalue=ABCG5

The order-orientation of the draft sequence was accomplished by

using:

Avid (<http://baboon.math.berkeley.edu/avid/>),Lagan (<http://lagan.stanford.edu/>) and paired end information.

Funding agent: Programs for Genomic Applications (NHLBI)

Summary Statistics:

Sequencing vector: Plasmid, pUC18

Chemistry: Dye-terminator Big Dye

Assembly Program: Phrap version 0.990329.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* of the gaps between them are given, however the sizes

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

CC The present invention describes an activation-induced cytidine deaminase
CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
CC cytidine activity similar to APOBEC-1. AID has anti-allergic, anti-inflammatory,
CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
CC and can be used in gene therapy. AID polynucleotides are useful in
CC methods for identifying drugs for the treatment of B cell associated
CC immune system disorders, immunodeficiency diseases and allergies, such as
CC immunoglobulin A (IgA) deficiency disease, IGA nephritis, gamma-
CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
CC drug allergy, allergic rhinitis, Rosen disease, DiGeorge disease, ataxia
CC telangiectasia, common variable immunodeficiency disorder, MHC (major
CC histocompatibility class II deficiency disease, AIDS (auto
CC immunodeficiency syndrome), elevated IGE deficiency disorder and Igg subclass
CC selection disorder. The DNA sequences encoding AID may be used for gene
CC therapy and the antibodies to the AID protein may be used for diagnosis
CC and treatment of these disorders. The present sequence represents a
CC genomic DNA sequence of human AID
XX

Sequence 6564 BP; 1909 A; 1358 C; 1383 G; 1914 T; 0 U; 0 Other;

Query Match 100.0%; Score 300; DB 3; Length 6564;

Best Local Similarity 100.0%; Pred. No. 2.3e-63;

Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGGGGCCTGTAATCCAGCTACTCAGAGAGCTGAGGCGAGAGATCCCGGAGCTGGCA 60

1 GGGGGCCTGTAATCCAGCTACTCAGAGAGCTGAGGCGAGAGATCCCGGAGCTGGCA 60

61 GATCGGCTGAGGCTGGAGGTTGAGGCTACAGTAAAGATGATGAGTATGATTC 120

61 GATCGGCTGAGGCTGGAGGTTGAGGCTACAGTAAAGATGATGAGTATGATTC 120

121 AGCGTGGCGACAAAGTGAACCGTAAACAAAAAATTTAAAAAAGAAATTAG 180

121 AGCGTGGCGACAAAGTGAACCGTAAACAAAAAATTTAAAAAAGAAATTAG 180

181 ATCAAGATCCAACTGTAATAAGTGGCTTAAACCAATTAAGATTGAGTTATTC 240

181 ATCAAGATCCAACTGTAATAAGTGGCTTAAACCAATTAAGATTGAGTTATTC 240

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

28-MAR-2000; 2000W0-JP001918.

XX 29-MAR-1999; 99GP-00087192.
PR 24-JUN-1999; 99GP-00178999.
PR 27-DEC-1999; 99GP-00371382.

XX (NIB) JAPAN TOBACCO INC.
PA (HONU/) HONJO T.

PI Honjo T, Muramatsu M,
XX WPI; 2000-611715/58.

DR Nucleic acid encoding activation induced cytidine deaminase, useful as a
XX target for drug development for immune-related diseases including
PT allergies.

PT target for drug development for immune-related diseases including
PT allergies.

PS Claim 17; Page 163-170; 174pp; Japanese.

PS Claim 17; Page 163-170; 174pp; Japanese.

CC The present invention describes an activation-induced cytidine deaminase
CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
CC cytidine activity similar to APOBEC-1. AID has anti-allergic, anti-inflammatory,
CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
CC and can be used in gene therapy. AID polynucleotides are useful in
CC methods for identifying drugs for the treatment of B cell associated
CC immune system disorders, immunodeficiency diseases and allergies, such as
CC immunoglobulin A (IgA) deficiency disease, IGA nephritis, gamma-
CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
CC drug allergy, allergic rhinitis, Rosen disease, DiGeorge disease, ataxia
CC telangiectasia, common variable immunodeficiency disorder, MHC (major
CC histocompatibility class II deficiency disease, AIDS (auto
CC immunodeficiency syndrome), elevated IGE deficiency disorder and Igg subclass
CC selection disorder. The DNA sequences encoding AID may be used for gene
CC therapy and the antibodies to the AID protein may be used for diagnosis
CC and treatment of these disorders. The present sequence represents a
CC genomic DNA sequence of human AID
XX

Sequence 11204 BP; 3305 A; 2273 C; 2373 G; 3253 T; 0 U; 0 Other;

Query Match 99.5%; Score 298.4; DB 3; Length 11204;

Best Local Similarity 99.7%; Pred. No. 6.5e-63;

Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GGGGGCCTGTAATCCAGCTACTCAGAGAGCTGAGGCGAGAGATCCCGGAGCTGGCA 60

5216 GGGGGCCTGTAATCCAGCTACTCAGAGAGCTGAGGCGAGAGATCCCGGAGCTGGCA 5275

61 GATCGGCTGAGGCTGGAGGTTGAGGCTACAGTAAAGATGATGAGTATGATTC 120

5276 GATCGGCTGAGGCTGGAGGTTGAGGCTACAGTAAAGATGATGAGTATGATTC 5335

121 AGCGTGGCGACAAAGTGAACCGTAAACAAAAAATTTAAAAAAGAAATTAG 180

5336 AGCGTGGCGACAAAGTGAACCGTAAACAAAAAATTTAAAAAAGAAATTAG 5395

181 ATCAAGATCCAACTGTAATAAGTGGCTTAAACCAATTAAGATTGAGTTATTC 240

5396 ATCAAGATCCAACTGTAATAAGTGGCTTAAACCAATTAAGATTGAGTTATTC 5455

241 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 300

5456 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 5515

5456 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 5515

5456 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 5515

5456 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 5515

5456 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 5515

5456 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 5515

5456 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 5515

5456 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 5515

5456 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 5515

5456 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 5515

5456 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 5515

5456 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 5515

5456 TGAGGCGAAGAGAACCATCAGGGGCTTTCAGCATGGGAATGGCATGTGACCTGGT 5515

04-DEC-2002 (first entry)
DNA encoding human translocation del(12p) protein #1.

KM Chromosome aberration; oncogenic fusion protein; cancer;
KM Proliferative disease; cellular protein isoform; heat shock protein 90;
KM HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
KM T cell lymphoma; B cell lymphoma; chronic myeloid leukemia; CML;
KM acute myeloid leukemia; AML; chronic myelomonocytic leukemia; CMML;
KM papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
KM rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200269900-A2.
XX
PD 12-SEP-2002.
XX
PF 01-MAR-2002; 2002WO-US006518.
XX
PR 01-MAR-2001; 2001US-0272751P.
XX
PA (CONF-) CONFORMA THERAPEUTICS CORP.
XX
PI Fritz LC, Burrows FJ;
XX
DR WPI; 2002-698710/75.
XX
DR P-PSDB; ABG95082.
XX
PT Treating genetically-defined disease associated with chromosomal
PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
PT diseases, involves administering an inhibitor of heat shock protein 90.
XX
PS Disclosure; Page 242-245; 389pp; English.
XX
CC The invention describes a method of treating genetically-defined disease
CC associated with chromosomal aberrations yielding oncogenic fusion
CC proteins (I), treating cancerous cells containing (I) in a heterogeneous
CC cell population, treating proliferative diseases associated with mutant
CC protein or cellular protein isoform (II) dependent on heat shock protein
CC (HSP)-90, or selectively treating cells expressing (II) involving
CC administering HSP90-inhibitor. The method is useful for treating
CC genetically-defined disease with chromosomal aberration yielding
CC oncogenic fusion protein, treating cancerous cells containing fusion
CC protein in heterogeneous cell population, treating proliferative disease
CC (e.g. Rheumatoid arthritis or cancer) associated with mutant protein or
CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
CC p53), or selectively treating cells expressing mutant protein or cellular
CC protein isoform in a patient heterozygous for (II). The method is useful
CC for treating a disease e.g. haematopoietic disorder such as T or B cell
CC lymphoma, chronic myeloid leukemia (CML), APL, ALL, NHL and CMML,
CC or a disease characterised by a solid tumour such as papillary thyroid
CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
CC synovial sarcoma. The method is also useful for treating viral
CC infections. This represents the DNA sequence of a chromosome aberration
XX
SQ Sequence 11204 BP; 3305 A; 2273 C; 2373 G; 3253 T; 0 U; 0 Other;
Query Match 99.5%; Score 298.4; DB 6; Length 11204;
Best Local Similarity 99.7%; Pred. No. 6; se-63;
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGGGCCTGTATATCCAGCTACTCAGAGGCTGAGGAGATCCGGAGCCTGGCA 60
Db 5216 GGGGCGCTGTATATCCAGCTACTCAGAGGCTGAGGAGATCCGGAGCCTGGCA 5275
QY 61 GATCTGCTTGAAGCTGGAGAGTTGAGGCTACAGTAAGCCAGATCATGATCATCTTC 120
Db 5276 GATCTGCTTGAAGCTGGAGAGTTGAGGCTACAGTAAGCCAGATCATGATCATCTTC 5335
QY 121 AGCCTGGGCGCAAAAGTGAAGACCGTAAACAATAAATTTAAATAAAGAAATTAG 180
Db 5336 AGCCTGGGCGCAAAAGTGAAGACCGTAAACAATAAATTTAAATAAAGAAATTAG 5395
QY 181 ATCAAGATCCAACTGTAAAAAGTGGCCTTAACACCATTTAAAGAGTTTGAATTATTC 240

Db 5396 ATCAAGATCCAACTGTAAAAAGTGGCCTTAACACCATTTAAAGAGTTTGAATTATTC 5455
QY 241 TGGAGGCGAAGAGAGACATCAGGAGGCTTCAAGCATGGGATATGGATGCACTGTGT 300
Db 5456 TGGAGGCGAAGAGAGACATCAGGAGGCTTCAAGCATGGGATATGGATGCACTGTGT 5515
RESULT 4
AAL50814/C
ID AAL50814 standard; DNA; 79528 BP.
XX
AC AAL50814;
XX
DT 30-JAN-2003 (first entry)
XX
DE Human cancer status prediction method-related DNA sequence #6.
XX
KW Human; gene therapy; cancer status prediction; cancer; ds;
KW cancer malignancy evaluation; drug design; antisense nucleic acid.
XX
OS Homo sapiens.
XX
PN WO200272828-A1.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-JP002153.
XX
PR 14-MAR-2001; 2001JP-00073063.
XX
PR 06-APR-2001; 2001JP-00108503.
XX
PR 02-AUG-2001; 2001JP-00234807.
XX
PA (DNAC-) DNA CHIP RES INC.
PA (HISF) HITACHI SOFTWARE ENG CO LTD.
XX
PI Kato K, Iwao K, Noguchi S, Matoba R;
XX
DR WPI; 2002-713517/77.
XX
PT Computer-aided statistical method for predicting cancer, applicable in
PT gene therapy for evaluating cancer malignancy with data for use in drug
PT design.
XX
PS Disclosure; Page 84-131; 182pp; Japanese.
XX
CC The invention comprises a method for predicting cancer status. The method
CC involves: measuring expression doses of genes obtained from specimens;
CC selecting at least one gene as the gene for an assay; using the
CC measurement results on expression doses of the selected genes for
CC multivariate analysis; and classifying the specimens in analogous groups
CC with results of the multivariate analysis on expression patterns of the
CC genes. The method of the invention is useful for predicting cancer, which
CC is applicable in gene therapy for evaluating cancer malignancy with data
CC for use in drug design (e.g. antisense nucleic acids for use in gene
CC therapy to treat cancer). The present DNA sequence represents a human
CC nucleic acid of the invention
XX
SQ Sequence 79528 BP; 19015 A; 20270 C; 20468 G; 19775 T; 0 U; 0 Other;
Query Match 36.4%; Score 109.2; DB 6; Length 79528;
Best Local Similarity 72.7%; Pred. No. 1; se-16;
Matches 141; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 1 GGGGGCCTGTATATCCAGCTACTCAGAGGCTGAGGAGATCCGGAGCCTGGCA 60
Db 64094 GGGGCGCTGTATATCCAGCTACTCAGAGGCTGAGGAGATCCGGAGCCTGGCA 64035
QY 61 GATCTGCTTGAAGCTGGAGAGTTGAGGCTACAGTAAGCCAGATCATGATCATCTTC 120
Db 64094 GATCTGCTTGAAGCTGGAGAGTTGAGGCTACAGTAAGCCAGATCATGATCATCTTC 63975
QY 121 AGCCTGGGCGCAAAAGTGAAGACCGTAAACAATAAATTTAAATAAAGAAATTAG 180

Db 63974 AGCGTGGAGACAGAAAGACTCCATCTCAAAAAAAAAAATAATAATAATAATAATAATAA 63911

Qy 181 ATCAAGATCCACT 194
| | | | |
Db 63914 ATAATTAAATTAAT 63901

RESULT 5
ABQ88117
ID ABQ88117 standard; cDNA; 129722 BP.
XX
XX ABQ88117;
XX
XX 18-SEP-2002 (first entry)
DE Human osteoblast differentiation related CDNA SEQ ID NO 24.
KW Human; osteoblast; stem cell differentiation; bone tissue deposition;
KW osteoporosis; osteopathic; ss.
XX
XX Homo sapiens.
OS
XX WO200250301-A2.
PN
XX 27-JUN-2002.
PD
XX 18-DEC-2001; 2001WO-US048276.
PF
XX 18-DEC-2000; 2000US-0255882P.
PR 24-APR-2001; 2001US-0285691P.
XX
PA (GENE-) GENE LOGIC INC.
XX (PROC) PROCTER & GAMBLE CO.
PI
PI Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
PI Mertz L;
DR WPL; 2002-557663/59.

PT Use of genes and their expression profiles associated with osteoblast
PT differentiation for screening modulators bone formation, for diagnosing
PT or treating e.g. osteoporosis, or as markers for the differentiation
PT process.

Claim 1; SEQ ID NO 24; 78bp + Sequence Listing; English.

XX The invention relates to genes and their expression profiles are used
CC for: (a) screening modulators of precursor stem cell differentiation into
CC osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
CC deposition of bone tissue, abnormal rate of osteoblast formation or
CC osteoporosis; or (c) treating or monitoring treatment of the conditions
CC cited in (b); or monitoring the progression of bone tissue deposition.
CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
CC induced abnormalities in bone formation or bone loss, conditions that
CC involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
CC or fibrous dysplasia. The present sequence is that of an osteoblast
CC differentiation associated CDNA marker of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences

Qy Query Match 32.7%; Score 99; DB 6; Length 129722;
Best Local Similarity 76.4%; Pred. No. 8.1e-14;
Matches 133; Conservative 0; Mismatches 40; Indels 1; Gaps 1;

1 GGGGCGCCTGTATCCAGTACTGAGAGGCTGAGAGCGAGAGTGCCGGAAGCTTGACA 60
49787 GGAGCGCTGTATATCCAGTACTCGGAGGCTGAGAGCGAGTAGTCGTGAGGTCAG -A 49844

| Qy | 61 | GATCTGCTGAGCTGGAGGATTGAGGCTACAGTAAGCCAAAGATCATGCGCATATCTTC | 120 |
|----------|---|--|-------|
| Db | 49946 | GATATGGCTTAACCTTGGAGAGCAGAGTTTGCACTGAGCCGAGATTGTGCCACTGCACCTCC | 49905 |
| Qy | 121 | AGCCTGGGCGCAACAAGTAGAGCCGTAAACAAAAAAAAAATTTTAAAAAGAA | 174 |
| Db | 49906 | AGCCTGGGTACACAGAGTGAACCTTGTCTCAAAAAAAAAAAAAAAAAAGAA | 49959 |
| RESULT 6 | | | |
| ID | AA03615 | standard; DNA; 32134 BP. | |
| AC | AA03615; | | |
| DT | 21-NOV-2001 | (first entry) | |
| DE | Human reproductive system related antigen DNA SEQ ID NO: 6303. | | |
| XX | Human; reproductive system related antigen; reproductive system disorder; | | |
| KM | Human; gene therapy; ds. | | |
| XX | Homo sapiens. | | |
| OS | MO200155320-A2. | | |
| PN | MO200155320-A2. | | |
| XX | 02-AUG-2001. | | |
| PF | 17-JAN-2001; 2001MO-US001339. | | |
| XX | | | |
| PR | 31-JAN-2000; | 2000US-0179065P. | |
| PR | 04-FEB-2000; | 2000US-0180628P. | |
| PR | 24-FEB-2000; | 2000US-0184664P. | |
| PR | 02-MAR-2000; | 2000US-0186350P. | |
| PR | 16-MAR-2000; | 2000US-0189874P. | |
| PR | 17-MAR-2000; | 2000US-0190076P. | |
| PR | 18-APR-2000; | 2000US-0198123P. | |
| PR | 19-MAY-2000; | 2000US-0205515P. | |
| PR | 07-JUN-2000; | 2000US-0209467P. | |
| PR | 28-JUN-2000; | 2000US-0214886P. | |
| PR | 30-JUN-2000; | 2000US-0215135P. | |
| PR | 07-JUL-2000; | 2000US-0216647P. | |
| PR | 07-JUL-2000; | 2000US-0216880P. | |
| PR | 11-JUL-2000; | 2000US-0217487P. | |
| PR | 11-JUL-2000; | 2000US-0217496P. | |
| PR | 14-JUL-2000; | 2000US-0218290P. | |
| PR | 26-JUL-2000; | 2000US-0220963P. | |
| PR | 26-JUL-2000; | 2000US-0220964P. | |
| PR | 14-AUG-2000; | 2000US-0224518P. | |
| PR | 14-AUG-2000; | 2000US-0224519P. | |
| PR | 14-AUG-2000; | 2000US-0225213P. | |
| PR | 14-AUG-2000; | 2000US-0225214P. | |
| PR | 14-AUG-2000; | 2000US-0225266P. | |
| PR | 14-AUG-2000; | 2000US-0225267P. | |
| PR | 14-AUG-2000; | 2000US-0225268P. | |
| PR | 14-AUG-2000; | 2000US-0225270P. | |
| PR | 14-AUG-2000; | 2000US-0225477P. | |
| PR | 14-AUG-2000; | 2000US-0225757P. | |
| PR | 14-AUG-2000; | 2000US-0225758P. | |
| PR | 14-AUG-2000; | 2000US-0225759P. | |
| PR | 18-AUG-2000; | 2000US-0226279P. | |
| PR | 22-AUG-2000; | 2000US-0226861P. | |
| PR | 22-AUG-2000; | 2000US-0226868P. | |
| PR | 22-AUG-2000; | 2000US-0227182P. | |
| PR | 23-AUG-2000; | 2000US-0227009P. | |
| PR | 30-AUG-2000; | 2000US-0228924P. | |
| PR | 01-SEP-2000; | 2000US-0228287P. | |
| PR | 01-SEP-2000; | 2000US-0228343P. | |
| PR | 01-SEP-2000; | 2000US-0228344P. | |
| PR | 01-SEP-2000; | 2000US-0228345P. | |
| PR | 05-SEP-2000; | 2000US-0228503P. | |
| PR | 05-SEP-2000; | 2000US-0228513P. | |
| PR | 06-SEP-2000; | 2000US-0230437P. | |

[illegible]

| | | |
|---------------------------|--|---|
| PR | 17-NOV-2000; | 2000US-0249244P. |
| PR | 17-NOV-2000; | 2000US-0249245P. |
| PR | 17-NOV-2000; | 2000US-0249264P. |
| PR | 17-NOV-2000; | 2000US-0249265P. |
| PR | 17-NOV-2000; | 2000US-0249297P. |
| PR | 17-NOV-2000; | 2000US-0249299P. |
| PR | 17-NOV-2000; | 2000US-0249300P. |
| PR | 01-DEC-2000; | 2000US-0250160P. |
| PR | 01-DEC-2000; | 2000US-0250391P. |
| PR | 05-DEC-2000; | 2000US-0251030P. |
| PR | 05-DEC-2000; | 2000US-0251988P. |
| PR | 05-DEC-2000; | 2000US-0256719P. |
| PR | 06-DEC-2000; | 2000US-0251479P. |
| PR | 08-DEC-2000; | 2000US-0251856P. |
| PR | 08-DEC-2000; | 2000US-0251868P. |
| PR | 08-DEC-2000; | 2000US-0251869P. |
| PR | 08-DEC-2000; | 2000US-0251989P. |
| PR | 08-DEC-2000; | 2000US-0251990P. |
| PR | 11-DEC-2000; | 2000US-0254097P. |
| PR | 05-JAN-2001; | 2001US-0259678P. |
| PA | (HUMA-) HUMAN GENOME SCI INC. | |
| PI | Rosen CA, Barash SC, Ruben SM; | |
| XX | WPI; 2001-465570/50. | |
| XX | Isolated nucleic acid molecule encoding a reproductive system antigen is | |
| PT | used in preventing, treating or ameliorating a medical condition. | |
| XX | Disclosure; SEQ ID NO 6303; 1297bp + Sequence Listing; English. | |
| PS | The present invention provides the protein and coding sequences of a | |
| CC | number of human reproductive system related antigens. These can be used | |
| CC | in the prevention and treatment of reproductive system disorders, | |
| CC | including cancer. The present sequence is a genomic sequence encoding a | |
| CC | protein of the invention | |
| XX | Sequence 32134 BP; 9171 A; 6268 C; 6931 G; 9764 T; 0 U; 0 Other; | |
| SQ | | |
| Query Match | 31.3%; | Score 94; DB 4; Length 32134; |
| Best Local Similarity | 71.3%; | Pred. No. 5.5e-13; |
| Matches 124; Conservative | 0; Mismatches | 50; Indels 0; Gaps 0; |
| QY | 1 | GGGGGCGCTGAATCCCGACGTACTCAGGAGGCTGAGCGCAGAGGANTCCGGGAGCTTGGA 60 |
| DB | 12455 | GGGCACTGTACTCTCCACGCTACTCTGGGAGGCTGAGCGCAGAGGATGGCATGAGGAGGA 12514 |
| QY | 61 | GATCTGCTTGAAGCCTGGAGGTTGAGGCTACAGTAAAGCCAGATCATGCCAGTATCTTC 120 |
| DB | 12515 | GAATGGCATGAACCCAGAGGAGGAGGCTTGCACTGAGCCGAGATCACACCATGTGACATCC 12574 |
| QY | 121 | AGCTTGGGCGCAAAAGTGAGACCGCTAACAAAAAATTTAAAAAGAA 174 |
| DB | 12575 | AGCCTGGGTGACAAAGTAGAGATCCCTCTCAAAAAAATTTAAAAATATA 12628 |
| RESULT 7 | | |
| ID | ABA07813 | standard; DNA; 32134 BP. |
| XX | AC | ABA07813; |
| XX | DT | 11-JAN-2002 (first entry) |
| DE | Human ovarian and breast cancer associated polynucleotide SEQ ID NO 608. | |
| XX | Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral; | |
| KW | antiallergic; hepatotropic; antidiabetic; antiinflammatory; antilicer; | |
| KW | vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic; | |
| KW | cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder; | |
| XX | neurological disease; infection; human; secreted protein; ds. | |

OS Homo sapiens.
XX
XX WC200155325-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US001345.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246522P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249248P.
PR 17-NOV-2000; 2000US-0249249P.
PR 17-NOV-2000; 2000US-0249250P.
PR 17-NOV-2000; 2000US-0249251P.
PR 17-NOV-2000; 2000US-0249252P.
PR 17-NOV-2000; 2000US-0249253P.
PR 17-NOV-2000; 2000US-0249254P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249256P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249258P.
PR 17-NOV-2000; 2000US-0249259P.
PR 17-NOV-2000; 2000US-0249260P.
PR 17-NOV-2000; 2000US-0249261P.
PR 17-NOV-2000; 2000US-0249262P.
PR 17-NOV-2000; 2000US-0249263P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249266P.
PR 17-NOV-2000; 2000US-0249267P.
PR 17-NOV-2000; 2000US-0249268P.
PR 17-NOV-2000; 2000US-0249269P.
PR 17-NOV-2000; 2000US-0249270P.
PR 17-NOV-2000; 2000US-0249271P.
PR 17-NOV-2000; 2000US-0249272P.
PR 17-NOV-2000; 2000US-0249273P.
PR 17-NOV-2000; 2000US-0249274P.
PR 17-NOV-2000; 2000US-0249275P.
PR 17-NOV-2000; 2000US-0249276P.
PR 17-NOV-2000; 2000US-0249277P.
PR 17-NOV-2000; 2000US-0249278P.
PR 17-NOV-2000; 2000US-0249279P.
PR 17-NOV-2000; 2000US-0249280P.
PR 17-NOV-2000; 2000US-0249281P.
PR 17-NOV-2000; 2000US-0249282P.
PR 17-NOV-2000; 2000US-0249283P.
PR 17-NOV-2000; 2000US-0249284P.
PR 17-NOV-2000; 2000US-0249285P.
PR 17-NOV-2000; 2000US-0249286P.
PR 17-NOV-2000; 2000US-0249287P.
PR 17-NOV-2000; 2000US-0249288P.
PR 17-NOV-2000; 2000US-0249289P.
PR 17-NOV-2000; 2000US-0249290P.
PR 17-NOV-2000; 2000US-0249291P.
PR 17-NOV-2000; 2000US-0249292P.
PR 17-NOV-2000; 2000US-0249293P.
PR 17-NOV-2000; 2000US-0249294P.
PR 17-NOV-2000; 2000US-0249295P.
PR 17-NOV-2000; 2000US-0249296P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249298P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251038P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251858P.
PR 08-DEC-2000; 2000US-0251866P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX
XX
PA (HUMA-) HUMAN GENOME SCT INC.

| | | |
|----|-------------|-----------------|
| PR | 14-AUS-2000 | 2000US-0225447P |
| PR | 14-AUS-2000 | 2000US-0225575P |
| PR | 14-AUS-2000 | 2000US-0225758P |
| PR | 14-AUS-2000 | 2000US-0225759P |
| PR | 18-AUS-2000 | 2000US-0226279P |
| PR | 22-AUS-2000 | 2000US-0226681P |
| PR | 22-AUS-2000 | 2000US-0226686P |
| PR | 22-AUS-2000 | 2000US-0227182P |
| PR | 23-AUS-2000 | 2000US-0228092P |
| PR | 30-AUS-2000 | 2000US-0228924P |
| PR | 01-SEP-2000 | 2000US-0229287P |
| PR | 01-SEP-2000 | 2000US-0229343P |
| PR | 01-SEP-2000 | 2000US-0229344P |
| PR | 01-SEP-2000 | 2000US-0229345P |
| PR | 05-SEP-2000 | 2000US-0229509P |
| PR | 05-SEP-2000 | 2000US-0229513P |
| PR | 06-SEP-2000 | 2000US-0230437P |
| PR | 06-SEP-2000 | 2000US-0230438P |
| PR | 08-SEP-2000 | 2000US-0231242P |
| PR | 08-SEP-2000 | 2000US-0231243P |
| PR | 08-SEP-2000 | 2000US-0231244P |
| PR | 08-SEP-2000 | 2000US-0231433P |
| PR | 08-SEP-2000 | 2000US-0231434P |
| PR | 08-SEP-2000 | 2000US-0231441P |
| PR | 08-SEP-2000 | 2000US-0232080P |
| PR | 12-SEP-2000 | 2000US-0231968P |
| PR | 14-SEP-2000 | 2000US-0232397P |
| PR | 14-SEP-2000 | 2000US-0232398P |
| PR | 14-SEP-2000 | 2000US-0232399P |
| PR | 14-SEP-2000 | 2000US-0232400P |
| PR | 14-SEP-2000 | 2000US-0232401P |
| PR | 14-SEP-2000 | 2000US-0233063P |
| PR | 14-SEP-2000 | 2000US-0233064P |
| PR | 14-SEP-2000 | 2000US-0233065P |
| PR | 21-SEP-2000 | 2000US-0234223P |
| PR | 21-SEP-2000 | 2000US-0234274P |
| PR | 25-SEP-2000 | 2000US-0234977P |
| PR | 25-SEP-2000 | 2000US-0234988P |
| PR | 26-SEP-2000 | 2000US-0235484P |
| PR | 27-SEP-2000 | 2000US-0235834P |
| PR | 29-SEP-2000 | 2000US-0235837P |
| PR | 29-SEP-2000 | 2000US-0236372P |
| PR | 29-SEP-2000 | 2000US-0236373P |
| PR | 29-SEP-2000 | 2000US-0236374P |
| PR | 29-SEP-2000 | 2000US-0236375P |
| PR | 29-SEP-2000 | 2000US-0236376P |
| PR | 29-SEP-2000 | 2000US-0236377P |
| PR | 29-SEP-2000 | 2000US-0236378P |
| PR | 29-SEP-2000 | 2000US-0236379P |
| PR | 29-SEP-2000 | 2000US-0236380P |
| PR | 29-SEP-2000 | 2000US-0236381P |
| PR | 29-SEP-2000 | 2000US-0236382P |
| PR | 29-SEP-2000 | 2000US-0236383P |
| PR | 29-SEP-2000 | 2000US-0236384P |
| PR | 29-SEP-2000 | 2000US-0236385P |
| PR | 29-SEP-2000 | 2000US-0236386P |
| PR | 29-SEP-2000 | 2000US-0236387P |
| PR | 29-SEP-2000 | 2000US-0236388P |
| PR | 29-SEP-2000 | 2000US-0236389P |
| PR | 29-SEP-2000 | 2000US-0236390P |
| PR | 29-SEP-2000 | 2000US-0236391P |
| PR | 29-SEP-2000 | 2000US-0236392P |
| PR | 29-SEP-2000 | 2000US-0236393P |
| PR | 29-SEP-2000 | 2000US-0236394P |
| PR | 29-SEP-2000 | 2000US-0236395P |
| PR | 29-SEP-2000 | 2000US-0236396P |
| PR | 29-SEP-2000 | 2000US-0236397P |
| PR | 29-SEP-2000 | 2000US-0236398P |
| PR | 29-SEP-2000 | 2000US-0236399P |
| PR | 29-SEP-2000 | 2000US-0236400P |
| PR | 29-SEP-2000 | 2000US-0236401P |
| PR | 29-SEP-2000 | 2000US-0236402P |
| PR | 29-SEP-2000 | 2000US-0236403P |
| PR | 29-SEP-2000 | 2000US-0236404P |
| PR | 29-SEP-2000 | 2000US-0236405P |
| PR | 29-SEP-2000 | 2000US-0236406P |
| PR | 29-SEP-2000 | 2000US-0236407P |
| PR | 29-SEP-2000 | 2000US-0236408P |
| PR | 29-SEP-2000 | 2000US-0236409P |
| PR | 29-SEP-2000 | 2000US-0236410P |
| PR | 29-SEP-2000 | 2000US-0236411P |
| PR | 29-SEP-2000 | 2000US-0236412P |
| PR | 29-SEP-2000 | 2000US-0236413P |
| PR | 29-SEP-2000 | 2000US-0236414P |
| PR | 29-SEP-2000 | 2000US-0236415P |
| PR | 29-SEP-2000 | 2000US-0236416P |
| PR | 29-SEP-2000 | 2000US-0236417P |
| PR | 29-SEP-2000 | 2000US-0236418P |
| PR | 29-SEP-2000 | 2000US-0236419P |
| PR | 29-SEP-2000 | 2000US-0236420P |
| PR | 29-SEP-2000 | 2000US-0236421P |
| PR | 29-SEP-2000 | 2000US-0236422P |
| PR | 29-SEP-2000 | 2000US-0236423P |
| PR | 29-SEP-2000 | 2000US-0236424P |
| PR | 29-SEP-2000 | 2000US-0236425P |
| PR | | |

| Query Match | 31.3% | Score 94 | DB 4 | Length 32191 |
|-----------------------|--------------|-------------------|---------------|--------------|
| Best Local Similarity | 71.3% | Pred. No. 5.5e-13 | | |
| Matches 124 | Conservative | 0 | Mismatches 50 | Indels 0 |
| | | | Gaps | 0 |

PT and prognostic assays for eating disorders, such as anorexia and bulimia nervosa.

XX Disclosure; Page 94-107; 149pp; English.

XX The invention relates to a novel isolated nucleic acid molecule comprising a variant gene associated with an eating disorder and selected from any of 119 polymorphisms with their corresponding genotyping in CC dataset, alleles and XBASE identification, given in the specification.

CC The novel nucleic acid molecule has polymorphisms in the serotonin receptor 1D, delta-opioid receptor, or dopamine receptor D2, which is useful in diagnostic and prognostic assays for eating disorders, in CC particular anorexia nervosa and bulimia nervosa. This polynucleotide CC sequence represents a human serotonin receptor 1D gene of the invention XX

SQ Sequence 24420 BP; 6212 A; 5915 C; 5886 G; 6407 T; 0 U; 0 Other;

Query Match 31.2%; Score 93.6; DB 7; Length 24420;
Best Local Similarity 71.5%; Pred. No. 6.4e-11;
Matches 123; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 3 GGGCCTGTAATCCAGCTACTCAGAGGCTGAGCAGAGATCCGGAGCCTGGCAGA 62
DB 17317 GTGCCTACAGTCCAGCTACTCGGAGGCTGAAGTGGAGAGCACCCTGAGTGGAGG 17258
QY 63 TCTGCTGAGCCTGGAGGTTGAGGCTACAGTAAGCCAAAGATCATGCCAGTATCTTCA 122
DB 17257 TCCACCTGAGTCTGGAGGTCAGGCGCGCAGTGAAGTATGTGCACCTGCACTCCAG 17198
QY 123 CCTGGGCGCAAGTGAAGACCGCTAACAAAAAATTTTAAAAAAGA 174
DB 17197 CCTAGGTGACAGAGCAGTACCCTGCTCAAAAAAGAAACAAAGAAAA 17146

RESULT 12
AAC93353
ID AAC93353 standard; cDNA; 1877 BP.
XX AAC93353;
XX
XX 16-FEB-2001 (first entry)
XX
XX Human secreted protein cDNA sequence #44.
XX
XX Cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
XX valnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein; ss.
XX
XX Homo sapiens.
XX
XX WO200058495-A1.
XX
XX 05-OCT-2000.
XX
XX 23-MAR-2000; 2000WO-US007661.
XX
XX 26-MAR-1999; 99US-0126504P.
XX
XX 07-JAN-2000; 2000US-0174847P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsoulis G;
XX
XX WPI; 2000-611720/58.
XX
XX P-PSDB; AAB51422.
XX
XX New nucleic acid molecules encoding 45 human secreted proteins for
XX diagnosing, preventing, treating or ameliorating medical conditions and
XX used as food additives or preservatives.
XX
XX Claim 1; Page 352; 410pp; English.

XX The invention relate to the isolation of genes AAC93310-C93354 encoding CC 45 human secreted proteins AAB51380-B51423. The genes can be used to CC generate fusion proteins by linking to the gene for the human CC immunoglobulin G Fc portion (AAC93301) for increasing the stability of CC the fusion protein as compared to the human protein only. The genes and CC proteins are useful for preventing, ameliorating or treating medical CC conditions, e.g. by protein or gene therapy. The genes are isolated from CC a range of human tissues disclosed in the specification. The nucleic CC acids, proteins, antibodies and (ant)agonists are useful in the CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and CC ovarian cancer, and other cancers of the adrenal gland, bone, bone CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) CC immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing CC; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) CC infectious diseases such as viral, bacterial, fungal and parasitic CC

SQ Sequence 1877 BP; 543 A; 358 C; 372 G; 596 T; 0 U; 8 Other;

Query Match 31.1%; Score 93.4; DB 3; Length 1877;
Best Local Similarity 72.5%; Pred. No. 3.8e-13;
Matches 121; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 3 GGGCCTGTAATCCAGCTACTCAGAGGCTGAGCAGAGATCCGGAGCCTGGCAGA 62
DB 1706 GGGCCTGTAATCCAGCTACTCAGAGGCTGAGCAGAGATCCGGAGCCTGGTGA 1765
QY 63 TCTGCTGAGCCTGGAGGTTGAGGCTACAGTAAGCCAAAGATCATGCCAGTATCTTCA 122
DB 1766 ATTGCTTGAACTCCGGAGGCGGAGGTTGAGTGAAGCCAGATGATGCACTGCACTCCAG 1825
QY 123 CCTGGGCGCAAGTGAAGACCGCTAACAAAAAATTTTAAAA 169
DB 1826 TCTAGGCAACAGGCGCAAACTCCATCTCAAAAAAATTTTAAAA 1872

RESULT 13
AAL38336/C
ID AAL38336 standard; DNA; 143899 BP.
XX AAL38336;
XX
XX 15-AUG-2002 (first entry)
XX
XX Genomic sequence encoding a human NGR2 protein.
XX
XX Cerebroprotective; neuroprotective; cytoskeletal; Nogo receptor homologue;
XX NGR2; NGR3; axonal growth; central nervous system; CNS; cerebral injury;
XX spinal cord injury; stroke; demyelinating disease; multiple sclerosis;
XX monophasic demyelination; encephalomyelitis; Marchiafava-Bignami disease;
XX multifocal leukoencephalopathy; panencephalitis; Spongy degeneration;
XX Alexander's disease; Canavan's disease; metachromatic leukodystrophy;
XX Krabbe's disease; immune; bait protein; genetic mapping; gene therapy;
XX transgenic animal; unregulated cellular growth; cancer; tumour; human;
XX gene; ds.
XX
XX Homo sapiens.
XX
XX WO200229059-A2.
XX
XX 11-APR-2002.
XX
XX 06-OCT-2001; 2001WO-US031488.
XX
XX 06-OCT-2000; 2000US-0238361P.
XX
XX (UYTA) UNIV YALE.
XX
XX (BIOI) BIOGEN INC.

PI Scitlatter SM, Cate RJ, Sah DWY;
 DR WPI; 2002-416677/44.
 XX
 XX Novel Nogo receptor homolog polypeptide, NGR2 or NGR3, useful for
 PT treating central nervous system disorder, cerebral injury, spinal cord
 PT injury, stroke, and demyelinating diseases.
 XX

PS Example 2; Page 176-214; 277pp; English.

XX The invention relates to a Nogo receptor homologue polypeptide, NGR2 or
 CC NGR3, comprising a 50 amino acid LRRT sequence, a 284 amino acid NTLRRCT
 CC sequence, or a 420, 461 or 392 amino acid sequence, all given in the
 CC specification. The NGR3 protein or its binding antibody is useful for
 CC decreasing inhibition of axonal growth of a central nervous system (CNS)
 CC neuron, by contacting the neuron NGR3 or its antibody, and for treating
 CC CNS disease, disorder or injury. NGR3 or a vector comprising NGR3 is
 CC useful for treating cerebral injury, spinal cord injury, stroke,
 CC demyelinating diseases, e.g. multiple sclerosis, monophasic
 CC demyelination, encephalomyelitis, multifocal leukoencephalopathy,
 CC paraneoplastic Marchiafava-Bignami disease, spongy degeneration,
 CC Alexander's disease, Canavan's disease, metachromatic leukodystrophy and
 CC Krabbe's disease. NGR3 is useful for inducing an immune response in a
 CC mammal against NGR3, as a bait protein in a two-hybrid or three-hybrid
 CC assay, and as a research tool for identification, characterization and
 CC purification of interacting, regulatory proteins. The nucleotide
 CC sequences of the invention are useful for screening for RFP associated
 CC with certain disorders, for genetic mapping, and for gene therapy. The
 CC vector containing NGR3 is useful for producing non-human transgenic
 CC animals. The NGR3 binding antibody is useful for isolating and purifying
 CC NGR3, for localisation and/or quantitation of NGR3, and for diagnostic
 CC and therapeutic purposes. The sequences of the invention, vectors and
 CC antibodies are useful for treating or preventing unregulated cellular
 CC growth such as cancer and tumour growth. This polynucleotide sequence
 CC represents the genomic sequence encoding a human NGR2 protein of the
 CC invention
 CC
 CC
 CC

SQ Sequence 143899 BP; 36346 A; 35277 C; 35318 G; 35657 T; 0 U; 1301 Other;

Query Match 31.1%; Score 93.2; DB 6; Length 143899;
 Best Local Similarity 71.8%; Pred. No. 1.2e-12;
 Matches 122; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY CCTGTAATCCAGTCTACTCAGAGGCTGAGGACAGAGAGATCCGGGAGCTGCGAATCT 65
 DB 3935 CTTTAGTCTTCAAGCTACTTGGAGGCTGAGGATCACTTGGACATGGGAGGTC 3876
 QY GCCTGAGGCTGAGGAGTTGAGGCTACAGTAAGCAAGATCATGCGAGTAATCTCAGCT 125
 DB 3875 ACTTAGGCTTGAGAGTTGAGGCTGCGCTGAGCCATGATCAACCACTGCTCCAGTCT 3816
 QY 126 GGGCGACAAAGTGAAGCCGTAACAAAAAAAAAATTTAAAAAGAAA 175
 DB 3815 GGGTGACAGAGTGAAGTCTTAAATAAATTTAAAAAATCAAAA 3766

RESULT 14
 AAC04760/c
 ID AAC04760 standard; cDNA; 203 BP.

AC AAC04760;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 8835.

KM Human, 5' EST, expressed sequence tag, secreted protein; cDNA isolation;
 KM gene therapy; chromosome mapping; ss.

OS Homo sapiens.

XX Homo sapiens.
 XX EPI033401-A2.
 XX

PD 06-SEP-2000.

XX 21-FEB-2000; 2000EP-00200610.

PR 26-FEB-1999; 99US-0122487P.

PA (GSEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
 XX
 PS Claim 1; SEQ ID NO 8835; 71pp + Sequence listing; English.

XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
 CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
 CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
 CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors
 CC
 CC
 CC

SQ Sequence 203 BP; 29 A; 60 C; 42 G; 66 T; 0 U; 6 Other;

Query Match 30.7%; Score 92; DB 3; Length 203;
 Best Local Similarity 67.2%; Pred. No. 4.9e-13;
 Matches 121; Conservative 5; Mismatches 53; Indels 0; Gaps 0;

QY 2 GGGGCTGTAATCCAGTCTACTCAGAGGCTGAGGACAGAGAGATCCGGGAGCTGCGAG 61
 DB 183 GGTGCTGTAATCCAGTCTACTCAGAGGCTGAGGACAGAGAGATCCGGGAGCTGCGAG 124
 QY 62 ATCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 121
 DB 123 ATGAGCTGTAATCCAGTCTACTCAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 64
 QY 122 GCCTGGGCGACAAAGTGAAGCCGTAACAAAAAAAAAATTTAAAAAGAAAATTTAGA 181
 DB 63 GCCTGGGCGACAGAGCGAGATCCCTTAAATAAATTTAAAAAATTAATAAAGTATA 4

RESULT 15
 ABS69895/c
 ID ABS69895 standard; DNA; 56737 BP.

AC ABS69895;

DT 21-NOV-2002 (first entry)

DE Human hypoxanthine-guanine phosphoribosyltransferase (HPRT) gene.

KM Vector; adenovirus; adeno-associated; adenosine deaminase gene; receptor;
 KM adenosine deaminase deficiency; severe combined immune deficiency; PAH;
 KM beta-chain; haemoglobin gene; beta-thalassaemia; sickle cell disease;
 KM low density lipoprotein gene; familial hypercholesterolaemia;
 KM hypoxanthine-guanine phosphoribosyltransferase; Lesch-Nyhan syndrome;
 KM phenylalanine hydroxylase gene; gene therapy; phenylketonuria;
 KM dystrophin gene; muscular dystrophy; cystic fibrosis; immunosuppressant;
 KM human cystic fibrosis transmembrane conductance regulator gene;
 KM antihaemic; antilipemic; noctropic; cytosaric; dermatological; human;
 KM alpha-1-antitrypsin; lysosomal glucocerebrosidase; ADA; HPRT;
 KM lysosomal arylsulphatase A; ornithine transcarbamylase; ARSA; OTC; NF;

KM purin nucleoside phosphorylase; gene; ds.

OS Homo sapiens.

PN US2002102731-A1.

PD 01-AUG-2002.

PF 12-FEB-2001; 2001US-00782378.

PR 02-OCT-2000; 2000US-0237747P.

PA (UYNV) UNIV NEW YORK STATE RES FOUND.

PI Hearing P, Bahou WF, Sandalon Z, Gatenko DV;

DR WPI; 2002-690619/74.

PT Producing vector, by introducing vector having nucleotide sequence,
PT adenovirus inverted terminal repeats and packaging sequence, and adeno-
PT associated virus terminal repeat, into cell, and culturing cell.

PS Disclosure; Page 93-117; 191pp; English.

CC The present invention relates to a new method of producing a vector. The
CC method involves introducing recombinant vector having nucleotide sequence
CC (NS) having 5' and 3' end, left and right inverted terminal repeats of
CC adenovirus flanking NS, adenovirus packaging sequence linked to inverted
CC terminal repeat, and adeno-associated virus terminal repeat linked to 3'
CC end of NS, into cell expressing adenovirus early gene lacking from vector
CC ; and culturing cell to produce another vector. The method is useful for
CC generating vectors, especially mad vectors. The method is useful in
CC transferring nucleotide sequences of interest into a cell, for gene
CC transfer applications (e.g. gene therapy) in vitro, ex vivo and in vivo.
CC The nucleotide sequences are useful for treating diseases associated with
CC 1c, i.e. adenosine deaminase gene associated with adenosine deaminase
CC deficiency with severe combined immune deficiency, beta-chain of
CC haemoglobin gene associated with beta-thalassemia and sickle cell
CC disease, receptor for low density lipoprotein gene associated with
CC familial hypercholesterolemia, hypoxanthine-guanine
CC phosphoribosyltransferase associated with Lesch-Nyhan syndrome,
CC phenylalanine hydroxylase (PAH) gene associated with phenylketonuria,
CC dystrophin gene associated with muscular dystrophy, and human cystic
CC fibrosis transmembrane conductance regulator gene associated with cystic
CC fibrosis. The present nucleic acid sequence represents a human disease
CC gene sequence that was used in the methods of the invention

SQ Sequence 56737 BP; 15689 A; 11281 C; 11599 G; 18168 T; 0 U; 0 Other;

Query Match 30.3%; Score 91; DB 6; Length 56737;

Best Local Similarity 65.9%; Pred. No. 3,4e-12; Mismatches 1; Gaps 1;

Matches 147; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

QY 1 GGGGGCCTGTATCCAGCTACTCAAGAGGCTGAGGCAAGAGATCCGCGAGCTTGGC 59

DB 3504 GCGCACTGTATGTCCTGCTACTTGGAGGCTGAAGTGGAGAAATCCCTGAGCTGGGG 3445

QY 60 AGATCTGCTGAGCTTGGAGGTTGAGGCTACAGTAAGCCAAAGATCATGCCAGTATCTT 119

DB 3444 AGATCACTCCGAGCCCGAGGAGTCAAGCTGCAAGTCAAGTATGTCCTGCTGCACTC 3385

QY 120 CAGCTGGGCGAACAAGTGAACCGTAACAAAAAAATTTAAAAAAGAAATTTA 179

DB 3384 CATCTAGGTGACAGATGAACCCCTGTCTCAAAAAAAGAAATTTGCGCAATTTAAGTAA 3325

QY 180 GATCAAGATCCAACTGTAAAAAGTGGCTTAAACACCAATTA 222

DB 3324 GTTCATGTTTGAAGATGAAAAATCAACATTTTTCCTCAGCA 3282


```

: FILE REFERENCE: CLO00164CIP
: CURRENT APPLICATION NUMBER: US/09/916,204
: CURRENT FILING DATE: 2001-07-24
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 48763
: TYPE: DNA
: ORGANISM: Human
: US-09-916-204-3

```

| | | | | |
|-----------------------|-----------------|--------------------|------------|---------------|
| Query Match | 27.2%; | Score 81.6; | DB 4; | Length 48763; |
| Best Local Similarity | 68.5%; | Pred. No. 4.1e-12; | | |
| Matches 137; | Conservative 0; | Mismatches 44; | Indels 19; | Gaps 1; |

| | | |
|----|--|-------|
| Oy | 2GGGGCCCTTATATCCAGCTACTCAGAGAGGCTAGGCGAGAGATCCGGGAGACCTGGCAG | 61 |
| Db | 16610 GGCACCTTATATCCAGCTACTCAGAGAGGCTAGGCGAGAGAT----- | 16567 |

62 ATCTGCTCCGAGCCCTGGGAGGTTGAGGCTACACATAAGCAAGATCATGCCAGATATCTTCA 121
 Oy
 1656 ---TgCTTGAACTCAGGAGTGAAGTTGCAATGAGGCAAGATCCGACCACTGCACTCCA 16510
 Db

QY 122 GCGCTGGGCGACAAAGTGAAGCCCTAACAAAAAAAAAAAATTAAAAAGAAATTGCA 181

Db 16509 GCGTGGGTGACAGACAGCAAGCTTACCTAAAAAATTTTTTTTTTTTTTTTTTTT 16450

```

QY      182 TCAAGATCCAACCTGTAAAA 201
          ||| ||| ||| ||| |||
Db      16449 GAAAAAACCAAAAAAAA 16430

```

RESULT 3
US-09-328-925-4
Sequence 4. Application US/09328925

```

; Patent No. 6610906
; GENERAL INFORMATION:
; APPLICANT: Kurachi, Kotoku
; APPLICANT: Kurachi, Sumiko
; TITLE OF INVENTION: Nucleotide Sequences for Gene Regulation and Methods of
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: UM-03603
; CURRENT APPLICATION NUMBER: US/09/328,925
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ. ID NOS.: 84
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO. 4
; LENGTH: 38059
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-328-925-4

```

| | | | | |
|-----------------------|--------------|--------------------|----------------|---------------|
| Query Match | 27.1%; | Score 81.4; | DB 4; | Length 38059; |
| Best Local Similarity | 65.8%; | Pred. No. 4.3e-12; | | |
| Matches 144; | Conservative | 0; | Mismatches 56; | Indels 19; |
| | | | | Gaps 1; |

| | | | |
|----|------|---|-----|
| Oy | 1 | GGGGGCTGTAAATCCAGCTACTACGAGAGGCTGAGGCAGGAGATCCGCGAGCTGGCA | 60 |
| | | GGGTCCTGTAAATCCAGCTACTCTTGGAGAGCTGAGGCAGGAGAAAT | 750 |
| Db | 7457 | GGGTCCTGTAAATCCAGCTACTCTTGGAGAGCTGAGGCAGGAGAAAT | |

| | | | |
|----|------|--|-----|
| QY | 61 | GATCGCCTGAGCCCTGGAGAGTTGAGGCTACAGTAAAGCCAGATCATGCCAGATAACTTC | 120 |
| Db | 7503 | -----GCTTGAACCTGGAGAGTTGAGAGTTGAGTGAAGCTGATGCGCACCACTGCACTCC | 755 |

| | | | | | | |
|----|------|----------------------------|---------|-----------|------------------|-----|
| Qy | 121 | AGCCCTGGCGCAAACTGAGACCGTAA | CAAAAAA | TTTTAAAAA | GAATTTAG | 180 |
| Db | 7558 | AGCCTGGCGCAGAGCAAGACTCTCT | CAAAAAA | AAAAAGAGT | TGAAGAAAAAGTCTAG | 761 |

Dy 181 ATCAGATCCAACTGTAAAAAGTGGCCCAAAACCCCAT 219
||| ||| ||| |||
Db 7618 GCTAATTCAAAGAAAAAGTAGGCCCAAAAAGAATT 7656

RESULT 4
US-09-426-290-1/c
Convergence 1 Application RUC/09426290

Patent No. 6410712
GENERAL INFORMATION:
APPLICANT: Berglind Ran Olafsdottir
APPLICANT: Jeffery Gujcher

TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
FILE REFERENCE: 2345.2001-000
CURRENT APPLICATION NUMBER: US/09/426, 290
CURRENT FILING DATE: 1999-10-25

```

NUMBER OF SEQ ID NOS: 24
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 168575

```

TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS

```
LOCATION: (21181)...(21405)
NAME/KEY: CDS
LOCATION: (95252)...(95430)
NAME/KEY: CDS
```

```
LOCATION: (101753) ... (101958)
NAME/KEY: CDS
LOCATION: (110324) ... (110439)
NAME/KEY: CDS
```

```

LOCATION: (127058) ... (127278)
NAME/KEY: CDS
LOCATION: (127009) ... (127130)
NAME/KEY: CDS

```

DESCRIPTION: (120210) ... (120210)
 IS-09-426-290-1
 Query Match 27.1%;

Matches 130; Conservative 0; Mismatches 33; Indels 19; Gaps 1

8856 8857 8858 8859 8860 8861 8862 8863 8864 8865 8866 8867 8868 8869 8870 8871 8872 8873 8874 8875 8876 8877 8878 8879 8880 8881 8882 8883 8884 8885 8886 8887 8888 8889 8890 8891 8892 8893 8894 8895 8896 8897 8898 8899 8900 8901 8902 8903 8904 8905 8906 8907 8908 8909 8910 8911 8912 8913 8914 8915 8916 8917 8918 8919 8920 8921 8922 8923 8924 8925 8926 8927 8928 8929 8930 8931 8932 8933 8934 8935 8936 8937 8938 8939 8940 8941 8942 8943 8944 8945 8946 8947 8948 8949 8950 8951 8952 8953 8954 8955 8956 8957 8958 8959 8960 8961 8962 8963 8964 8965 8966 8967 8968 8969 8970 8971 8972 8973 8974 8975 8976 8977 8978 8979 8980 8981 8982 8983 8984 8985 8986 8987 8988 8989 8990 8991 8992 8993 8994 8995 8996 8997 8998 8999 9000

88566 ----GCTTGAAACCGGAGGTGAGGCTGCAGTGAAGCCGAGATTTCACCACTGCACCTCC 88571

121 AGCGTGGCGACAAAGTGAGACCGTAAACAAAAAATTTAAAAAAGAAATTTAG 180

88511 AGCTGGGCGACAGAGTGAAGGCTCCCTCTCAGAAAAAAAAAAAAAAAAATTG 0845

181 AT 182

88451 TT 88450

US-09-740-041-3/c
Sequence 3, Application US/09740041
Patent No. 6562593
CENTRAL INFORMATION.

APPLICANT: MERKULOV, Karl et al
TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS
TITLE OF INVENTION: AND USES THEREOF

FILE REFERENCE: CL001001
CURRENT APPLICATION NUMBER: US/09/740,041
CURRENT FILING DATE: 2000-12-20
NUMBER OF SEQ ID NOS: 4

```

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 66804
TYPE: DNA

```

ORGANISM: Human

| | | | |
|----|-------|-----------|-------|
| QY | 181 | ATCAAGATC | 189 |
| | | | |
| Db | 42296 | AGGTACATC | 42304 |

RESULT 8
US-09-618-166-79

? Sequence 79, Application US/05618106
 ? Patent No. 6583112
 ? GENERAL INFORMATION:
 ? APPLICANT: Fu, Ying-Hui
 ? Yu, Chang-Ea
 ? Oshima, Junko
 ? Mulligan, John T.
 ? Schellenberg, Gerald D.
 ? TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
 ? WERNER'S SYNDROME
 ? NUMBER OF SEQUENCES: 209
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESSEE: Seed Intellectual Property Law Group
 ? STREET: 701 Fifth Avenue, Suite 6300
 ? City: Seattle
 ? STATE: Washington
 ? COUNTRY: USA
 ? ZIP: 98104-7092
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: floppy disk
 ? COMPUTER: IBM PC compatible
 ? OPERATING SYSTEM: PC-DOS/MS-DOS
 ? SOFTWARE: Patent Release #1.0, Version #1.30
 ? CURRENT APPLICATION DATA:
 ? APPLICATION NUMBER: US/09/618,166
 ? FILING DATE: 17-Jul-2000
 ? CLASSIFICATION: <unknown>
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: Mcmasters, David D.
 ? REGISTRATION NUMBER: 33,963
 ? REFERENCE/DOCKET NUMBER: 240052.419C1
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: (206) 622-4900
 ? TELEFAX: (206) 682-6031
 ? INFORMATION FOR SEQ ID NO: 79:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 87350 base pairs
 ? TYPE: nucleic acid
 ? STRANDEDNESS: single
 ? TOPOLOGY: linear
 ? SEQUENCE DESCRIPTION: SEQ ID NO: 79:
 ? US-09-618-166-79

RESULT 9

```

US-09-791-211-3
Sequence 3, Application US/09791211
Patent No. 6448080
GENERAL INFORMATION:
APPLICANT: Donna T. Ward
APPLICANT: Andrew T. Walt
TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
FILE REFERENCE: RTS-0205
CURRENT APPLICATION NUMBER: US/09/791,211
CURRENT FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 90
SEQ ID NO 3
LENGTH: 87543
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: 7421
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 7427
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 11609
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 12605
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 12742
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29370
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29422
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 2979
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29980
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29981
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 30136
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 30140
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31205
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31206
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31592
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 33160
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 34066
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 34072
OTHER INFORMATION: unknown

```

```
/ NAME/KEY: unsure
/ LOCATION: 36816
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 39020
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 42164
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 42459
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 46808
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 46823
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 46826
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 47291
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 52785
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 52787
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 53384
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 54684
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 59215
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 59235
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 59242
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 63290
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 66614
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 68660
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 68697
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 68718
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 68733
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 68739
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 69785
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 79134
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/

/ LOCATION: 79198
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 86336
/ OTHER INFORMATION: unknown
/ OTHER INFORMATION:
/ US-09-791-211-3

Query Match      26.9%; Score 80.8; DB 4; Length 87543;
Best Local Similarity 69.8%; Pred. No. 8.2e-12;
Matches 132; Conservative 0; Mismatches 38; Indels 19; Gaps 1;

QY      1 GGGGGCCTGTAATCCAGCTACTCAGAGGCTGAGGAGGATCCGGAGCCTGGCA 60
Db      42328 GGGTGCTGTATATCCAGTATTCAGAGGCTGAGGAGGAGAT----- 42372
/
QY      61 GATTCGCTGAGGCTGGAGGTTGAGGCTACAGTAAGCCAGATCATGCCATATCTTC 120
Db      42373 ----TGCTTGAACCCAGAGGAGTGAGAGTTGCAGTGAATGTCACACTGCACCTTC 42428
/
QY      121 AGCCTGGGCGACAAAGTGAACCCGTAACAAATTTAAATTTAG 180
Db      42429 AGCCTGGGAGACAGAGCGAGACTCTGTCTCNAATTTAAATTTAG 42488
/
QY      181 ATCAAGATC 189
Db      42489 AGGTACATC 42497

RESULT 10
US-09-539-333D-1
/ Sequence 1, Application US/0953933D
/ Patent No. 6476208
/ GENERAL INFORMATION:
/ APPLICANT: Cohen, Daniel
/ APPLICANT: Blumenfeld, Marta
/ APPLICANT: Chumakov, Ilya
/ APPLICANT: Bougueterec, Lydie
/ APPLICANT: Bihaun, Bernard
/ APPLICANT: Beslioux, Laurent
/ TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLLELIC MARKERS
/ FILE REFERENCE: GENSET.047UUS
/ CURRENT APPLICATION NUMBER: US/09/539,333D
/ CURRENT FILING DATE: 2000-03-30
/ PRIOR APPLICATION NUMBER: US 60/126,903
/ PRIOR FILING DATE: 1999-03-30
/ PRIOR APPLICATION NUMBER: US 60/131,971
/ PRIOR FILING DATE: 1999-04-30
/ PRIOR APPLICATION NUMBER: US 60/132,065
/ PRIOR FILING DATE: 1999-04-30
/ PRIOR APPLICATION NUMBER: US 60/143,928
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: US 60/145,915
/ PRIOR FILING DATE: 1999-07-27
/ PRIOR APPLICATION NUMBER: US 60/146,453
/ PRIOR FILING DATE: 1999-07-29
/ PRIOR APPLICATION NUMBER: US 60/146,452
/ PRIOR FILING DATE: 1999-07-29
/ PRIOR APPLICATION NUMBER: US 60/162,288
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: US 09/416,384
/ PRIOR FILING DATE: 1999-10-12
/ NUMBER OF SEQ ID NOS: 231
/ SOFTWARE: Patent.pm
/ SEQ ID NO 1
/ LENGTH: 319608
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 31..1107
/ OTHER INFORMATION: 5'regulatory region g35018 gene
/ FEATURE:
```

NAME/KEY: exon
LOCATION: 1108..1289
OTHER INFORMATION: exon A g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 14877..14920
OTHER INFORMATION: exon B g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 18778..18862
OTHER INFORMATION: exon Bbis g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 25593..25740
OTHER INFORMATION: exon C g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 29388..29502
OTHER INFORMATION: exon D g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 29967..30282
OTHER INFORMATION: exon E g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 65505..65853
OTHER INFORMATION: exon G g35018 gene
FEATURE:
NAME/KEY: misc_feature
LOCATION: 65854..67854
OTHER INFORMATION: 3' regulatory region g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 94124..94964
OTHER INFORMATION: exon g35017
FEATURE:
NAME/KEY: exon
LOCATION: 201188..201234
OTHER INFORMATION: exon S g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 214676..214793
OTHER INFORMATION: exon T g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215702..215746
OTHER INFORMATION: exon U g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 216836..216915
OTHER INFORMATION: exon V g35030 gene
FEATURE:
NAME/KEY: misc_feature
LOCATION: 213818..215818
OTHER INFORMATION: 3' regulatory region g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215819..215941
OTHER INFORMATION: exon R complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215819..215975
OTHER INFORMATION: exon Rbis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 216661..216952
OTHER INFORMATION: exon Qbis complement g34872 gene
FEATURE:
NAME/KEY: exon

LOCATION: 216661..217061
OTHER INFORMATION: exon Q complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 217027..217061
OTHER INFORMATION: exon Q1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 229647..229742
OTHER INFORMATION: exon X complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 230408..230721
OTHER INFORMATION: exon P complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231272..231412
OTHER INFORMATION: exon Obis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231787..231880
OTHER INFORMATION: exon O2 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231870..231879
OTHER INFORMATION: exon O1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 234174..234321
OTHER INFORMATION: exon O complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 237406..237428
OTHER INFORMATION: exon Nbis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 239719..239807
OTHER INFORMATION: exon N2 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 239719..239853
OTHER INFORMATION: exon N complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240569
OTHER INFORMATION: exon M1117 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240596
OTHER INFORMATION: exon M1090 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240617
OTHER INFORMATION: exon M1069 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240644
OTHER INFORMATION: exon M2 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240934
OTHER INFORMATION: exon M692 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..241685
OTHER INFORMATION: exon M1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240800..240993

```

; OTHER INFORMATION: exon MS1 complement g34872 gene
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 241686..243685
; OTHER INFORMATION: 5' regulatory region g34872 gene
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 290652..292652
; OTHER INFORMATION: 3' regulatory region g34665 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 292653..292841

Query Match      26.9%; Score 80.8; DB 4; Length 319608;
Best Local Similarity 63.6%; Pred. No. 1.3e-11;
Matches 150; Conservative 0; Mismatches 67; Indels 19; Gaps 1;

QY      2 GGGGCGCTGTAATCCAGCTACTCAGAGAGCTGAGGAGAGATCCGCGAGCCTGGCAG 61
        |||
DB      191544 GGCACCTGTAAATCCAGCTACTCCGAGGCTGAGGAGAGATCA----- 191589
        |||

QY      62 ATCTGCGCTGAGCGCTGGAGTTGAGGCTACAGTACCAAGATCATGCTAGTACTCA 121
        |||
DB      191590 -----CTTGACCTGGAGATGAGAGTTGCAGTGCAGATCATGCTAGTACTCA 191644
        |||

QY      122 GCCTGGCGGACAAAGTGGAGCCGTACAAAAAATTTAAAAAAGAAATTTAGA 181
        |||
DB      191645 GCCTGGACAAACAAAGCAAGCTCTGTCTCAAAAAATTTAAAAATTTAAAAA 191704
        |||

QY      182 TCAAGATCCAAGTGTAAAAAGTGGCTTAACACCAATTAAGATTGAGTTA 237
        |||
DB      191705 AATAAAGGAATCCACAGATTGTATTCTTTAATACATTAAGTGTCTA 191760
        |||

RESULT 11
US-09-679-409-1
; Sequence 1, Application US/09679409
; Patent No. 655316
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumentfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bouguerelet, Lydie
; APPLICANT: Essioux, Laurent
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENE, PROTEINS AND BIALLELIC MARKERS
; FILE REFERENCE: 53.US15.CIP
; CURRENT FILING DATE: US/09/679,409
; PRIOR APPLICATION NUMBER: 2000-10-03
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 09/539,333
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 09/416,384
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 60/168,088
; PRIOR FILING DATE: 1999-11-30
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: Patent.Pm
; SEQ ID NO 1
; LENGTH: 319608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 199122..201122
; OTHER INFORMATION: 5' regulatory region
; NAME/KEY: exon
; LOCATION: 201123..201234
; OTHER INFORMATION: exon S
; NAME/KEY: exon
; LOCATION: 201123..201560
; OTHER INFORMATION: exon S2
; NAME/KEY: exon
; LOCATION: 214676..214793
; OTHER INFORMATION: exon T

; NAME/KEY: exon
; LOCATION: 215702..215746
; OTHER INFORMATION: exon U
; NAME/KEY: exon
; LOCATION: 216836..216984
; OTHER INFORMATION: exon V
; NAME/KEY: exon
; LOCATION: 216836..217077
; OTHER INFORMATION: exon V2
; NAME/KEY: exon
; LOCATION: 217671..217764
; OTHER INFORMATION: exon V1
; NAME/KEY: exon
; LOCATION: 227655..227736
; OTHER INFORMATION: exon V4
; NAME/KEY: exon
; LOCATION: 238715..238919
; OTHER INFORMATION: exon V3
; NAME/KEY: exon
; LOCATION: 240440..240673
; OTHER INFORMATION: exon W
; NAME/KEY: exon
; LOCATION: 240440..241153
; OTHER INFORMATION: exon W2
; NAME/KEY: exon
; LOCATION: 241072..241291
; OTHER INFORMATION: exon X
; NAME/KEY: exon
; LOCATION: 244353..244561
; OTHER INFORMATION: exon Y
; NAME/KEY: exon
; LOCATION: 246273..247802
; OTHER INFORMATION: exon Z
; NAME/KEY: misc.feature
; LOCATION: 247803..249803
; OTHER INFORMATION: 3' regulatory region
; NAME/KEY: allele
; LOCATION: 8316
; OTHER INFORMATION: 99-27943-150 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 21672
; OTHER INFORMATION: 99-27935-193 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 65485
; OTHER INFORMATION: 8-128-33 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 95396
; OTHER INFORMATION: 99-31960-363 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 107281
; OTHER INFORMATION: 99-24656-260 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 160640
; OTHER INFORMATION: 99-24639-163 : polymorphic base A or C
; NAME/KEY: allele
; LOCATION: 160876
; OTHER INFORMATION: 99-24634-108 : polymorphic base A or T
; NAME/KEY: allele
; LOCATION: 168974
; OTHER INFORMATION: 99-7652-162 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 170810
; OTHER INFORMATION: 99-16100-147 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 173358
; OTHER INFORMATION: 99-5862-167 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 189957
; OTHER INFORMATION: 99-5919-215 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 197163
; OTHER INFORMATION: 99-24658-410 : polymorphic base A or G
; NAME/KEY: allele
```

LOCATION: 200778 : polymorphic base A or G
OTHER INFORMATION: 8-303-235 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 202651 : polymorphic base A or G
OTHER INFORMATION: 8-300-221 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 202679 : polymorphic base A or G
OTHER INFORMATION: 8-300-193 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 203378 : polymorphic base A or T
OTHER INFORMATION: 8-299-128 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 204138 : polymorphic base A or T
OTHER INFORMATION: 8-296-213 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 204605 : polymorphic base C or T
OTHER INFORMATION: 8-252-190 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 204934 : polymorphic base A or G
OTHER INFORMATION: 99-24644-194 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 205206 : polymorphic base A or C
OTHER INFORMATION: 8-295-248 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 205329 : polymorphic base C or T
OTHER INFORMATION: 8-295-125 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 206064 : polymorphic base A or G
OTHER INFORMATION: 8-293-130 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 206545 : polymorphic base A or G
OTHER INFORMATION: 8-292-198 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 207313 : polymorphic base A or G
OTHER INFORMATION: 8-251-322 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 208285 : polymorphic base A or G
OTHER INFORMATION: 8-289-322 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 208960 : polymorphic base C or T
OTHER INFORMATION: 8-287-249 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 209123 : polymorphic base A or T
OTHER INFORMATION: 8-287-86 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 209631 : polymorphic base A or G
OTHER INFORMATION: 8-285-319 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 210361 : polymorphic base G or C
OTHER INFORMATION: 8-283-278 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210463 : polymorphic base A or G
OTHER INFORMATION: 8-283-176 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 210486 : polymorphic base G or C
OTHER INFORMATION: 8-283-153 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210583 : polymorphic base C or T
OTHER INFORMATION: 8-283-56 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 210879 : polymorphic base G or C
OTHER INFORMATION: 8-282-345 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210964 : polymorphic base G or T
OTHER INFORMATION: 8-282-260 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 210979 : polymorphic base A or C
OTHER INFORMATION: 8-282-245 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 211050 : variable motif AAAGG or GAAGGAAGGAAGGAAGA
OTHER INFORMATION: 8-282-174 : variable motif AAAGG or GAAGGAAGGAAGGAAGA
NAME/KEY: allele
LOCATION: 211132

OTHER INFORMATION: 8-282-92 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 211247 : polymorphic base A or G
OTHER INFORMATION: 8-281-367 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 211315 : polymorphic base A or G
OTHER INFORMATION: 8-281-299 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 211366 : polymorphic base G or C
OTHER INFORMATION: 8-281-248 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 212520 : polymorphic base A or C
OTHER INFORMATION: 8-279-197 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 212821 : polymorphic base C or T
OTHER INFORMATION: 8-278-289 : polymorphic base C or T
NAME/KEY: allele

Query Match 26.9%; Score 80.6; DB 4; Length 319608;
Best Local Similarity 63.6%; Pred. No. 1.3e-11;
Matches 150; Conservative 0; Mismatches 67; Indels 19; Gaps 1;

QY 2 GGGGCTGTAATCCAGCTACTCAGAGCTGAGGCAAGAGATCCGGAGCTGCAG 61
DB 191544 GGACCTGTAATACCACTACTCGGAGCTGAGGCAAGATCA----- 191589
QY 62 ATCTGCTGAGCTGAGGTTGAGCTACAGTAAGCAAGATCATGCCAGTACTTCA 121
DB 191590 -----CTGACCTGAGAGATGAGTTGCAAGTATGCGACTGACTTCA 191644
QY 122 GCTGGGCGACAAGTGAAGCCGTAACAAAAAAAAAATTAAAGAAATTGA 181
DB 191645 GCTGGACAACAAGCAAGACTCTGTCAAAAAAAAAATTAAATTAATAA 191704
QY 182 TCAGATCCACTGTAATAAAGTGGCTTAACCCACATTAAAGTTGAGTTA 237
DB 191705 AATTAAGTAATCCACAGATTGCTTATATTCTTAATACATGAATGATCTA 191760

RESULT 12
US-09-056-105-31
Sequence 31, Application US/09056105
Patent No. 6287569
GENERAL INFORMATION:
APPLICANT: KIPPS, THOMAS J.
APPLICANT: KU, YUNOJ
TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
FILE REFERENCE: 233/221
CURRENT APPLICATION NUMBER: US/09/056,105
CURRENT FILING DATE: 1998-04-06
EARLIER APPLICATION NUMBER: 60/043,467
EARLIER FILING DATE: 1997-04-10
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 31
LENGTH: 7130
TYPE: DNA
ORGANISM: Homo sapiens
US-09-056-105-31

Query Match 26.9%; Score 80.6; DB 3; Length 7130;
Best Local Similarity 66.0%; Pred. No. 3.9e-12;
Matches 142; Conservative 0; Mismatches 54; Indels 19; Gaps 1;
QY 5 GCGTGAATCCAGCTACTCAGAGCTGAGGCAAGAGATCCGGAGCTGCAGATC 64
DB 2755 GCGTGAATCCAGCTACTCAGAGCTGAGGCAAGAT----- 2795
QY 65 TGCTGAGCTGGAGGTTGAGGCTACAGTAAGCAAGTCAATGCCAGTACTTCAGCC 124
DB 2796 TGCTGAGCTGGAGGTTGAGGCTACAGTAAGCAAGTCAATGCCAGTACTTCAGCC 2855

| | | | |
|----|------|---|------|
| QY | 125 | TCGCGCAAAAGTGCAGCCGTACCAAAAAAAAAATTAAAAAAGAAATTTCATCA | 184 |
| Dd | 2856 | TGGGTGACAGAGTGTGACCTCCGCCCTCAAAAAAAAAAAAAAAAAAAAAAGTA | 2915 |
| QY | 185 | AGATCCAACTGTAAAAAGTGGCTTAAACCCACAT | 219 |
| Dd | 2916 | AGAAAAAGAAAAAGAAAGCAATTTTATCCCTGAT | 2950 |

```

RESULT 13
US-09-702-705-1626
; Sequence 1626, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darlick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Iqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1626
; LENGTH: 367
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-1626

```

| | Query Match | 26.8% | Score 80.4 | DB 4 | Length 367 |
|----|-----------------------|---|-------------------|---------------|------------|
| | Best Local Similarity | 67.1% | Pred. No. 1.6e-12 | | |
| | Matches 114 | Conservative | 0 | Mismatches 56 | Indels 0 |
| | | | | Gaps | 0 |
| QY | 5 | GCGCTTAATCCCAAGCTACTCAGAGAGCTGAGGACAGAGAGATCCGGAGCGCTTGCAGATC | 64 | | |
| Db | 153 | GCGCTTAATCCCAACTACTCAGAGAGCGGAGGACAGAGATTACTTGAACGACAGAGAT | 212 | | |
| QY | 65 | TGCGTCGAGCCTGGGAGGCTTTAGGCTTACAGTAAAGCCAAAGATCATGCGAGTATTCCTCAGCC | 124 | | |
| Db | 213 | CAGTCGACGCCCTGGGAGGAGGAGAGGTTGCAAGTGGCCGAGATTGCACACACTGTACTCCAGCC | 272 | | |
| QY | 125 | TGGGGGAGCAAAAGTGAGCCGTACACAAAATTTTAAAAAAGAA | 174 | | |
| Db | 273 | TGGGTGACAGACAAAGATCCATTCCTCAGTAATAATAATAATAATAATAA | 322 | | |

RESULT 14
 US-09-736-457-1626
 ; Sequence 1626, Application US/09736457
 ; Patent No. 6509448
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Lodes, Michael A.
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Vedwick, Tom
 ; APPLICANT: Carter, Darick
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Wang, Aijun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.478C15
 ; CURRENT APPLICATION NUMBERS: US/09/736,457
 ; CURRENT FILING DATE: 2000-12-13

```

; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1626
; LENGTH: 367
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-736-457-1626

```

| | Query Match | Score | DB # | Length |
|-----------------------|-------------|--------------|------|---------------------------------|
| Best Local Similarity | 67.18% | 26.8% | 80.4 | 367 |
| Matches | 114 | Conservative | 0 | Mismatches 56; Indels 0; Gaps 0 |

| | Query | Db | Score | DB # | Length |
|-----|--|-----|---|------|--------|
| 5 | GCCCTGATATCCAGCTACTCAGAGGCTGAGCGCAGAGAGATCCGGCGAGCCTGCGCATC | 153 | GCCCTGATATCCCACTACTCAGAAAGCGCGAGGCGAGAGATTACTTGAACCGACGAGAGAT | 212 | 64 |
| 65 | TGCCTGAGCCCTGGAGGTTGAGGGCTACAGTAGCCGACAGATCATGCGGTACTTACCTACGC | 213 | CATGCGAGCCCTGAGGAGGAGGTTGACGTAGGCCGAGATTGCACCACTGTATCTCCAGCC | 272 | 124 |
| 125 | TGGCGCAAAAGTGAGACCCGTAAACAAAAAATTTTAAAAAAGAA | 273 | TGGGTGACAGAGCAAGACTCCATCTCAGTAAATTTAAATTTAAAAA | 322 | 174 |

```

RESULT 15
US-09-614-124B-1626
; Sequence 1626, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jare
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.47869
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ. ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1626
; LENGTH: 367
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-614-124B-1626

```

| | Query Match | Similarity | Score | DB 4 | Length | 367 |
|----|------------------|---|-------------------|------|------------|-----|
| | Best Local Match | 67.1% | Pred. No. 1.6e-12 | | | |
| | Matches | 114 | Conservative | 0 | Mismatches | 56 |
| | | | | | Indels | 0 |
| | | | | | Gaps | 0 |
| Qy | 5 | GCTGTATATCCCACTACTCAGAGAGCTGAGGACAGAGATCCGCGAGCTGGCAGATC | 64 | | | |
| Db | 153 | GCTGTATATCCCACTACTCAGAGAGGCGAGGACAGAGAAATCTGTAACGACAGAGAT | 212 | | | |
| Qy | 65 | TGCTGAGCCTGGGAGGTTGAGGCTACAGTAAAGCCAAAGATATGCSAGTAACTTCAACC | 124 | | | |
| Db | 213 | CATGCGAGCCCTGGAGGCGAGAGGTTGAGTAGTGCCGAGATTGCACCACTGTACTCCAGCC | 272 | | | |
| Qy | 125 | TGGGCGCAAAAGTGAGAGCCGTAACAAAAAATTTAAAAAAGAA | 174 | | | |
| Db | 273 | TGGGTGACAGCAAGACTCAATTCAGTAAATATAATTAATTAATAA | 322 | | | |

Search completed: March 13, 2004, 00:51:01
Job time : 25.3214 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 21:22:51 ; Search time 105.276 Seconds
(without alignments)
10491.276 Million cell updates/sec

Title: US-09-966-880a-10_COPY_1_300
Perfect score: 300
Sequence: 1 99999ccctgaatccacgct.....aatgcatgtgcacatcgt 300

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 300 | 100.0 | 6564 | 9 | US-09-966-880a-10 |
| 2 | 298.4 | 99.5 | 11204 | 9 | US-09-966-880a-35 |
| 3 | 100 | 33.3 | 3523 | 15 | US-10-027-632-114084 |
| 4 | 94 | 31.3 | 32134 | 14 | US-09-764-891-6303 |
| 5 | 94 | 31.3 | 32134 | 14 | US-10-205-428-608 |
| 6 | 94 | 31.3 | 32191 | 10 | US-09-764-891-6304 |
| 7 | 94 | 31.3 | 32191 | 14 | US-10-205-428-609 |
| 8 | 93.4 | 31.1 | 669 | 15 | US-10-027-632-284884 |
| 9 | 93.2 | 31.1 | 143899 | 10 | US-09-972-546-15 |
| 10 | 91.4 | 30.5 | 1859 | 11 | US-09-984-429-549 |
| 11 | 91.4 | 30.5 | 12437 | 14 | US-10-314-321A-3 |
| 12 | 91 | 30.3 | 849 | 15 | US-10-027-632-29551 |
| 13 | 91 | 30.3 | 849 | 15 | US-10-027-632-29552 |
| 14 | 91 | 30.3 | 56737 | 9 | US-09-782-378A-17 |
| 15 | 90.6 | 30.2 | 595 | 15 | US-10-027-632-223820 |

| | | | | | | |
|------|------|------|-------|----|----------------------|-------------------|
| C 16 | 89.8 | 29.9 | 13893 | 13 | US-10-161-510-1 | Sequence 1, Appli |
| C 17 | 89 | 29.7 | 1236 | 15 | US-10-027-632-124822 | Sequence 124822, |
| C 18 | 89 | 29.7 | 1252 | 15 | US-10-027-632-257891 | Sequence 257891, |
| C 19 | 89 | 29.7 | 1252 | 15 | US-10-027-632-257892 | Sequence 257892, |
| C 20 | 89 | 29.7 | 1252 | 15 | US-10-027-632-257893 | Sequence 257893, |
| C 21 | 89 | 29.7 | 1252 | 15 | US-10-027-632-257894 | Sequence 257894, |
| C 22 | 89 | 29.7 | 1252 | 15 | US-10-027-632-257895 | Sequence 257895, |
| C 23 | 88.4 | 29.5 | 483 | 15 | US-10-027-632-215643 | Sequence 215643, |
| C 24 | 88 | 29.3 | 485 | 15 | US-10-027-632-21405 | Sequence 21405, A |
| C 25 | 88 | 29.3 | 485 | 15 | US-10-027-632-21406 | Sequence 21406, A |
| C 26 | 88 | 29.3 | 846 | 15 | US-10-027-632-160172 | Sequence 160172, |
| C 27 | 88 | 29.3 | 846 | 15 | US-10-027-632-160173 | Sequence 160173, |
| C 28 | 88 | 29.3 | 846 | 15 | US-10-027-632-160174 | Sequence 160174, |
| C 29 | 87.8 | 29.3 | 3157 | 15 | US-10-027-632-115461 | Sequence 115461, |
| C 30 | 87.8 | 29.3 | 6318 | 15 | US-10-247-335-1882 | Sequence 882, App |
| C 31 | 87.8 | 29.3 | 13069 | 9 | US-09-764-863-1850 | Sequence 1850, Ap |
| C 32 | 87.8 | 29.3 | 13069 | 14 | US-10-091-504-1850 | Sequence 1850, Ap |
| C 33 | 87.8 | 29.3 | 13069 | 15 | US-10-227-577-1850 | Sequence 1849, Ap |
| C 34 | 87.8 | 29.3 | 32203 | 9 | US-09-764-863-1849 | Sequence 1849, Ap |
| C 35 | 87.8 | 29.3 | 32203 | 14 | US-10-091-504-1849 | Sequence 1849, Ap |
| C 36 | 87.8 | 29.3 | 32203 | 15 | US-10-227-577-1849 | Sequence 1849, Ap |
| C 37 | 87.6 | 29.2 | 2445 | 15 | US-10-027-632-110185 | Sequence 110185, |
| C 38 | 87.4 | 29.1 | 668 | 15 | US-10-027-632-131147 | Sequence 131147, |
| C 39 | 86.6 | 28.9 | 473 | 15 | US-10-027-632-106241 | Sequence 106241, |
| C 40 | 86.4 | 28.8 | 2716 | 15 | US-10-027-632-111585 | Sequence 111585, |
| C 41 | 86.4 | 28.8 | 2716 | 15 | US-10-027-632-111586 | Sequence 111586, |
| C 42 | 86.2 | 28.7 | 482 | 15 | US-10-027-632-108468 | Sequence 108468, |
| C 43 | 86.2 | 28.7 | 482 | 15 | US-10-027-632-294298 | Sequence 294298, |
| C 44 | 86.2 | 28.7 | 482 | 15 | US-10-027-632-294298 | Sequence 294298, |
| C 45 | 85.8 | 28.6 | 497 | 10 | US-09-918-995-21543 | Sequence 21543, A |

ALIGNMENTS

| | |
|---|--|
| US-09-966-880a-10 | RESULT 1 |
| Sequence 10, Application US/09966880A | |
| Patent No. US20020164743A1 | |
| GENERAL INFORMATION: | |
| APPLICANT: Honjo, Tasuku | |
| APPLICANT: Muramatsu, Masamichi | |
| TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE | |
| FILE REFERENCE: 06501-088001 | |
| CURRENT APPLICATION NUMBER: US/09/966, 880A | |
| PRIOR FILING DATE: 2001-09-28 | |
| PRIOR APPLICATION NUMBER: PCT/JP00/01918 | |
| PRIOR FILING DATE: 2000-03-28 | |
| PRIOR APPLICATION NUMBER: JP 11-371382 | |
| PRIOR FILING DATE: 1999-12-27 | |
| PRIOR APPLICATION NUMBER: JP 11-178999 | |
| PRIOR FILING DATE: 1999-06-24 | |
| PRIOR APPLICATION NUMBER: JP 11-87192 | |
| PRIOR FILING DATE: 1999-03-29 | |
| NUMBER OF SEQ ID NOS: 36 | |
| SOFTWARE: FASTSEQ for Windows Version 4.0 | |
| SEQ ID NO 10 | |
| LENGTH: 6564 | |
| TYPE: DNA | |
| ORGANISM: Homo sapiens | |
| US-09-966-880a-10 | |
| Query Match | 100.0% ; Score 300 ; DB 9 ; Length 6564 ; |
| Best Local Similarity | 100.0% ; Pred. No. 1.2e-65 ; |
| Matches 300 ; Conservative 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ; | |
| QY 1 | GGGGGCTGTATATCCAGTACTACAGAGCTGAGGACAGAGAGATCCGGAGCTCGGA 60 |
| DB 1 | GGGGGCTGTATATCCAGTACTACAGAGCTGAGGACAGAGAGATCCGGAGCTCGGA 60 |
| QY 61 | GATTCGCTGAGGCTGGAGGTTGAGGCTACAGTACCAAGATCATGCTACATATCTTC 120 |
| DB 61 | GATTCGCTGAGGCTGGAGGTTGAGGCTACAGTACCAAGATCATGCTACATATCTTC 120 |

QY 121 AGCCTGGGCGACAAGTGAAGCCTTAACAAAAAATTTAAAAAAGAAATTTAG 180
DB 121 AGCCTGGGCGACAAGTGAAGCCTTAACAAAAAATTTAAAAAAGAAATTTAG 180
QY 181 ATCAAGATCCAACTGTATTAAGAGTGGCTTAACACCACTTAAGAGTTTGAAGTTATTC 240
DB 181 ATCAAGATCCAACTGTATTAAGAGTGGCTTAACACCACTTAAGAGTTTGAAGTTATTC 240
QY 241 TGCAGGCGAAGAGAACCAATCAAGGAGGCTTTCAGCATGGGAAATGGCATGGTGCACCTGGT 300
DB 241 TGCAGGCGAAGAGAACCAATCAAGGAGGCTTTCAGCATGGGAAATGGCATGGTGCACCTGGT 300

RESULT 2

US-09-966-880A-35
Sequence 35, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Matsumatsu, Masamichi
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966,880A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 35
LENGTH: 11204
TYPE: DNA
ORGANISM: Homo sapiens
US-09-966-880A-35

Query Match 99.5%; Score 298.4; DB 9; Length 11204;
Best Local Similarity 99.7%; Pred. No. 3,6e-65;
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGGGCTGTATATCCAGCTACTCAGAGGCTGAGGAGGAGATCCGCGAGCCTTGCA 60
DB 5216 GGGGGCTGTATATCCAGCTACTCAGAGGCTGAGGAGGAGATCCGCGAGCCTTGCA 5275
QY 61 GATTCGCTGAGCCTGGAGGTTGAGGCTACAGTAAGCCAGATATCCGCTATATTC 120
DB 5276 GATTCGCTGAGCCTGGAGGTTGAGGCTACAGTAAGCCAGATATCCGCTATATTC 5335
QY 121 AACCTGGGCGACAAGTGAAGCCTTAACACCACTTAAGAGTTTGAAGTTATTC 180
DB 5336 AACCTGGGCGACAAGTGAAGCCTTAACACCACTTAAGAGTTTGAAGTTATTC 5395
QY 181 ATCAAGATCCAACTGTATTAAGAGTGGCTTAACACCACTTAAGAGTTTGAAGTTATTC 240
DB 5396 ATCAAGATCCAACTGTATTAAGAGTGGCTTAACACCACTTAAGAGTTTGAAGTTATTC 5455
QY 241 TGCAGGCGAAGAGAACCAATCAAGGAGGCTTTCAGCATGGGAAATGGCATGGTGCACCTGGT 300
DB 5456 TGCAGGCGAAGAGAACCAATCAAGGAGGCTTTCAGCATGGGAAATGGCATGGTGCACCTGGT 5515

RESULT 3
US-10-027-632-114084
Sequence 114084, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

FILE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827,129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 114084
LENGTH: 3523
TYPE: DNA
ORGANISM: Human
US-10-027-632-114084

Query Match 33.3%; Score 100; DB 15; Length 3523;
Best Local Similarity 73.8%; Pred. No. 2.9e-15;
Matches 127; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 3 GGGGCTGTATATCCAGCTACTCAGAGGCTGAGGAGGAGATCCGCGAGCCTTGCA 62
DB 2365 GGGGCTGTATATCCAGCTACTCAGAGGCTGAGGAGGAGATCCGCGAGCCTTGCA 2424
QY 63 TGTGCTGAGCCTGGAGGTTGAGGCTACAGTAAGCCAGATATCCGCTATATTC 122
DB 2425 TGTGCTGAGCCTGGAGGTTGAGGCTACAGTAAGCCAGATATCCGCTATATTC 2484
QY 123 CCTGGGCGACAAGTGAAGCCTTAACAAAAAATTTAAAAAAGAA 174
DB 2485 CCTGGGCGACAAGTGAAGCCTTAACAAAAAATTTAAAAAAGAA 2536

RESULT 4
US-09-764-891-6303
Sequence 6303, Application US/09764891
Publication No. US20030077808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
PRIOR APPLICATION data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6303
LENGTH: 32134
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-891-6303

Query Match 31.3%; Score 94; DB 10; Length 32134;
Best Local Similarity 71.3%; Pred. No. 2.2e-13;
Matches 124; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 GGGGCTGTATATCCAGCTACTCAGAGGCTGAGGAGGAGATCCGCGAGCCTTGCA 60
DB 12455 GGGGCTGTATATCCAGCTACTCAGAGGCTGAGGAGGAGATCCGCGAGCCTTGCA 12514
QY 61 GATTCGCTGAGCCTGGAGGTTGAGGCTACAGTAAGCCAGATATCCGCTATATTC 120
DB 12515 GATTCGCTGAGCCTGGAGGTTGAGGCTACAGTAAGCCAGATATCCGCTATATTC 12574

QY 121 AGCTGGGCGCAAAAGTGAGACCGCTAACAAAAAATTTAAAAAGAA 174
DB 12575 AGCTGGGTGACAAAGTGAGACTCCGCTCAAAAAAATTTAAAAATAA 12628

RESULT 5

US-10-205-428-608
Sequence 608, Application US/10205428
Publication No. US20030108907A1
GENERAL INFORMATION:

APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P1117C1
CURRENT APPLICATION NUMBER: US/10/205,428
CURRENT FILING DATE: 2002-07-26

PRIOR APPLICATION NUMBER: 09/764,892
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1019
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 608
LENGTH: 32134
TYPE: DNA
ORGANISM: Homo sapiens
US-10-205-428-608

Query Match 31.3%; Score 94; DB 14; Length 32134;
Best Local Similarity 71.3%; Pred. No. 2.2e-13;
Matches 124; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 GGGGGCTGTATATCCAGCTACTAGAGAGCTGAGGAGAGAGATCCGGAGCTGGCA 60
DB 12455 GGGGACCTGTACTCCAGCTACTCGGAGGCTGAGGAGAGATGCGATGAGGAGGA 12514
QY 61 GATCTGCTGAGCTGGAGGTTGAGGCTACAGTAAGCAAGATCCAGTATCTTC 120
DB 12515 GAATGCGATGAAACCGAGGAGGAGGCTTGCAGTGAGCCGAGATCACACCATTCCTCC 12574
QY 121 AGCTGGGCGCAAAAGTGAGACCGCTAACAAAAAATTTAAAAAGAA 174
DB 12575 AGCTGGGTGACAAAGTGAGACTCCGCTCAAAAAAATTTAAAAATAA 12628

RESULT 6

US-09-764-891-6304
Sequence 6304, Application US/09764891
Publication No. US20030077808A1
GENERAL INFORMATION:

APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17

Prior Application data removed - consult PALM or file wrapper

NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6304
LENGTH: 32191
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-891-6304

Query Match 31.3%; Score 94; DB 10; Length 32191;
Best Local Similarity 71.3%; Pred. No. 2.2e-13;
Matches 124; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 GGGGGCTGTATATCCAGCTACTAGAGAGCTGAGGAGAGAGATCCGGAGCTGGCA 60
DB 9503 GGGGACCTGTACTCCAGCTACTCGGAGGCTGAGGAGAGATGCGATGAGGAGGA 9562
QY 61 GATCTGCTGAGCTGGAGGTTGAGGCTACAGTAAGCAAGATCCAGTATCTTC 120
DB 9563 GAATGCGATGAAACCGAGGAGGAGGCTTGCAGTGAGCCGAGATCACACCATTCCTCC 9622
QY 121 AGCTGGGCGCAAAAGTGAGACCGCTAACAAAAAATTTAAAAAGAA 174
DB 9623 AGCTGGGTGACAAAGTGAGACTCCGCTCAAAAAAATTTAAAAATAA 9676

RESULT 7

US-10-205-428-609
Sequence 609, Application US/10205428
Publication No. US20030108907A1
GENERAL INFORMATION:

APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P1117C1
CURRENT APPLICATION NUMBER: US/10/205,428
CURRENT FILING DATE: 2002-07-26
PRIOR APPLICATION NUMBER: 09/764,892
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1019
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 609
LENGTH: 32191
TYPE: DNA
ORGANISM: Homo sapiens
US-10-205-428-609

Query Match 31.3%; Score 94; DB 14; Length 32191;
Best Local Similarity 71.3%; Pred. No. 2.2e-13;
Matches 124; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 GGGGGCTGTATATCCAGCTACTAGAGAGCTGAGGAGAGAGATCCGGAGCTGGCA 60
DB 9503 GGGGACCTGTACTCCAGCTACTCGGAGGCTGAGGAGAGATGCGATGAGGAGGA 9562
QY 61 GATCTGCTGAGCTGGAGGTTGAGGCTACAGTAAGCAAGATCCAGTATCTTC 120


```
; Sequence 549, Application US/09984429
; Publication No. US20040010132A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 53 Human Secreted Proteins
; FILE REFERENCE: P2018P2
; CURRENT APPLICATION NUMBER: US/09/984,429
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,591
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/288,143
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: PCT/US98/21142
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/061,463
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,529
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/071,498
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,527
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,536
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,532
; NUMBER OF SEQ ID NOS: 727
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 549
; LENGTH: 1859
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-984-429-549

Query Match
Best Local Similarity 30.5%; Score 91.4; DB 11; Length 1859;
Matches 119; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 3 GGGCTGTAATCCAGCTACTCAGAGGCTGAGGAGGATCCGGAGCTTGAGAGA 62
DB 165 GCGGCTGTAGTCCAGGCTATTCAGAGGCTGAGAGGAGATTCCTGAACCGGTAGA 106
QY 63 TCTGCTGAGGCTGAGAGGTTGAGGCTACAGTAAGCAAGATCATGCGAGTATCTTCA 122
DB 105 ATTCTTGAATCCCGAGAGCGGAGGCTTGCACTGAGCCGAGATCTGCACTTCAG 46
QY 123 CCTGGGCGACAAAGTGAAGCCGTAACAAAAAAAAAAATTTAA 167
DB 45 TCTAGGCAACAGGCGCAAACTCCATCTCCAAAAAATTTAA 1

RESULT 11
US-10-314-321A-3
; Sequence 3, Application US/10314321A
; Publication No. US20030190648A1
; GENERAL INFORMATION:
; APPLICANT: Hitachi, Ltd.
; TITLE OF INVENTION: Gene Predicting Method
; FILE REFERENCE: 31010118US1
; CURRENT APPLICATION NUMBER: US/10/314,321A
; PRIOR FILING DATE: 2002-12-08
; PRIOR APPLICATION NUMBER: JP 2002-103333
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 12437
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-314-321A-3

Query Match
Best Local Similarity 30.5%; Score 91.4; DB 14; Length 12437;
Matches 72.1%; Pred. No. 7e-13;
```

```
Matches 119; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 5 GCGTGTAAATCCAGCTACTCAGAGGCTGAGGAGGATCCGGAGCTTGAGATC 64
DB 3122 GCGTGTAAATCCAGCTACTCAGAGGCTGAGGAGGATTCCTTGAGGATGAGAT 3181
QY 65 TCCCTGAGGCTGGAGGCTTGAAGCTACAGTAAAGCAATATGCAATATCTTCA 124
DB 3182 CTCTTGAAGCAGTGGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 3241
QY 125 TGGGGGCAAAAGTGAAGCCGTAACAAAAAAAAAAATTTAA 169
DB 3242 TGGGTGACAGAGCAAGCCATGTCTCAAAAAAACAACCA 3286

RESULT 12
US-10-027-632-29551
; Sequence 29551, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108627.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29551
; LENGTH: 849
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-29551

Query Match
Best Local Similarity 30.3%; Score 91; DB 15; Length 849;
Matches 121; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 4 GCGTGTAAATCCAGCTACTCAGAGGCTGAGGAGGATCCGGAGCTTGAGATC 63
DB 410 GCGTGTAAATCCAGCTACTCAGAGGCTGAGGAGGATTCGTTAGCCAGAGAGA 469
QY 64 CTGCTGAGGCTGGAGGCTTGAAGCTACAGTAAAGCAATATGCAATATCTTCA 123
DB 470 TCGCTTGAAGCCAGAGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 529
QY 124 CTGGGCGCAAAAGTGAAGCCGTAACAAAAAAAAAAATTTAA 174
DB 530 CTGGGTGAACAGAGCAAGCTTGTCCCAAAAAAAGAAAAA 580

RESULT 13
US-10-027-632-29552
; Sequence 29552, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108627.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29551
; LENGTH: 849
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-29552

Query Match
Best Local Similarity 30.3%; Score 91; DB 15; Length 849;
Matches 121; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
```

```

, FILE REFERENCE: 108827.129
, CURRENT APPLICATION NUMBER: US/10/027,632
, CURRENT FILING DATE: 2002-04-30
, PRIOR APPLICATION NUMBER: US 60/218,006
, PRIOR FILING DATE: 2000-07-12
, PRIOR APPLICATION NUMBER: US 60/198,676
, PRIOR FILING DATE: 2000-04-20
, PRIOR APPLICATION NUMBER: US 60/193,483
, PRIOR FILING DATE: 2000-03-29
, PRIOR APPLICATION NUMBER: US 60/185,218
, PRIOR FILING DATE: 2000-02-24
, PRIOR APPLICATION NUMBER: US 60/167,363
, PRIOR FILING DATE: 1999-11-23
, PRIOR APPLICATION NUMBER: US 60/156,358
, PRIOR FILING DATE: 1999-09-28
, PRIOR APPLICATION NUMBER: US 60/146,002
, PRIOR FILING DATE: 1999-08-09
, NUMBER OF SEQ ID NOS: 325720
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 29552
, LENGTH: 849
, TYPE: DNA
, ORGANISM: Human
, US-10-027-632-29552

```

| | | | | |
|-----------------------|--------------|--------------------|----------------|------------------|
| Query Match | 30.3%; | Score 91; | DB 15; | Length 849; |
| Best Local Similarity | 70.8%; | Pred. No. 3,2e-13; | | |
| Matches 121; | Conservative | 0; | Mismatches 50; | Indels 0; Gaps 0 |

| | | | |
|----|-----|---|-----|
| Qy | 4 | GGCCTGTATATCCAGCTACTCAGAGAGCTGAGGACAGAGATCCGCGAGCTTGCCAGAT | 63 |
| Db | 410 | GGCTTTGATGATCCAGGACTCCAGAGAGCTGAGGACAGAGATCGCTTTAGGCCACGAGAGA | 469 |
| Qy | 64 | CTGCGCTGAGCCTGGAGAGGTTGAGGCTACAGTAAGCCAAGATCATGCCAGTATATTAGC | 123 |
| Db | 470 | TCGCTTGAGGCCAGGAGGATTCAGCTTGACAGTATGATGTGGCCATCTGTACTCCAGC | 529 |
| Qy | 124 | CTGGGACACAAAGTGAACCGTAACTAAAAAAATTTAAAAAAAGAA | 174 |
| Db | 530 | CTGGGTACACAGACAAAGACTTTTCCCTCCAAAAAAGAAAAAGAAA | 560 |

```

RESULT 14
US-09-782-378A-17/C
Sequence 17, Application US/09782378A
Patent No. US20020102731A1
GENERAL INFORMATION:
APPLICANT: Hearing, Patrick
APPLICANT: Babou, Madie
APPLICANT: Sandelson, Ziv
APPLICANT: Gatenko, Dmitri
TITLE OF INVENTION: Amino-vital Vectors
FILE REFERENCE: STONW-04970
CURRENT APPLICATION NUMBER: US/09/782,378A
CURRENT FILING DATE: 2001-02-12
PRIOR APPLICATION NUMBER: 60/1237,747
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn version 3.0
SEQ ID NO 17
LENGTH: 56737
TYPE: DNA
ORGANISM: Homo sapiens
US-09-782-378A-17

```

| | Query Match | Similarity | Score | % | DB | Length | Mismatches | Indels | Gaps |
|--------------|-------------|--|---------|---|----|--------|------------|--------|------|
| Best Local | 147 | 65.9% | 1.5E-12 | | | | | | |
| Conservative | 0 | | | | | | 75 | 1 | 1 |
| QY | 1 | GGGGGCTGTATCCCACTACTCAGAGGCTGACGACGAGAGATCCCGGAGCCTGC | 59 | | | | | | |
| Db | 3504 | GCGACCTGTATGCCACGACTATTGGAGGCTGGAAGTGGAGAAATCCCTAGAGCTTGGG | 3445 | | | | | | |

| Accession | Sequence | Position |
|-----------|--|----------|
| Oy | AGATCTGCTCGAGGCTGGAGAGTTGAGGTACAGTAAGCCAAAGTCAATGCCAGTAATCTT | 119 |
| Oy | 60 | |
| Db | AGAAATCAACCCGAGCCCGGGAAGCTGAGGCTGCACTGAGCACTGATTTGTGGCACTGCACCTC | 3385 |
| Db | 3444 | |
| Oy | 120 | |
| Oy | CAGCCTGGGCGACAAAGTGAGACCGTAAACAAAAAATTTTAAAAAAGAAATTTA | 179 |
| Db | 3384 | |
| Db | CATCTTAGGTGACAGAGTAGAGACCCCTGCTCAAAAAAAGAAATTTGGCAGAAATTTAAGTA | 3325 |
| Oy | 180 | |
| Oy | GATCAAGATCCAACTGTAAAAAGTGGCCTAAACCAACATTTAA | 222 |
| Db | 3324 | |
| Db | GTGTAGTTTAAAGATGAAAAATCAATTTTTTCTTCAGCAA | 3282 |

```

RESULT 15
US-10-027-632-223820/c
; Sequence 223820, Application US/10027632
; Publication NO. US20030204075A5
; GENERAL INFORMATION:
; APPLICANT: Man'g, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ. ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 223820
; LENGTH: 595
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-223820

```

| | | | | |
|-----------------------|--------------|--|---------------|------------|
| Query Match | 30.2% | Score 90.6 | DB 15 | Length 595 |
| Best Local Similarity | 66.8% | Pred. No. 3,66-13 | | |
| Matches 129 | Conservative | 0 | Mismatches 64 | Indels 0 |
| Gaps | 0 | | | |
| QY | 3 | GAGCGCTGAATCCAGACTACTCAGAGGCTTAGCGCAGAGGATCCGCGAGACCTGGCAGA | 62 | |
| DB | 288 | GCGCCTTAATAGTCCAGCTACTTGAAGGCTAGGAGGAGATCACTGACCTGGGAGG | 229 | |
| QY | 63 | TCTGCGTAGCGCTGGAGGTGAGGCTTACAGTAAGCCAAATCATCCAGTATCTTAC | 122 | |
| DB | 228 | ATCATTTAGCCCAAGAGGTAGAGGCTTCAATGATGAGTATGATGGCATCACTGCAC | 169 | |
| QY | 123 | CCTGGGCGACAAAGTGAGACCGTAAACAAAAAAAAAAAAAAAAATTAAAGAAATTTAGAT | 182 | |
| DB | 168 | CCTGGGTGACAGAGTAGACCCCTGTCCCAAGCAACAAACAAAAAAAACAAAAGAA | 109 | |
| QY | 183 | CAAGATCCAACTG | 195 | |
| DB | 108 | ATCATTTGCAACAG | 96 | |

Search completed: March 13, 2004, 05:30:47
Job time : 107.276 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 13:19:34 ; Search time 955.276 Seconds
(without alignments)
9378.080 Million cell updates/sec

Title: US-09-966-880a-10_COPY_1_300
Perfect score: 300
Sequence: 1 GGGGAGCTGCTAATCCAGCT.....AATGCACTGCTGCTGCT 300

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
EST:
1: em_estdb:*
2: em_estdb:*
3: em_estdb:*
4: em_estdb:*
5: em_estdb:*
6: em_estdb:*
7: em_estdb:*
8: em_estdb:*
9: em_estdb:*
10: em_estdb:*
11: em_estdb:*
12: em_estdb:*
13: em_estdb:*
14: em_estdb:*
15: em_estdb:*
16: em_estdb:*
17: em_estdb:*
18: em_estdb:*
19: em_estdb:*
20: em_estdb:*
21: em_estdb:*
22: em_estdb:*
23: em_estdb:*
24: em_estdb:*
25: em_estdb:*
26: em_estdb:*
27: em_estdb:*
28: em_estdb:*
29: em_estdb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|-------|--------------------|--------|--------------------|
| 1 | 104.2 | 34.7 | 235 13 | BU665283 c1129b12 |
| 2 | 101.8 | 32.9 | 617 28 | AQ383054 RPT11-13 |
| 3 | 97.6 | 32.5 | 509 28 | AQ035617 CIT-HSP-2 |
| 4 | 97 | 32.3 | 526 28 | AQ51557 RPT1-11-3 |

ALIGNMENTS

| Result 1 | LOCUS | DEFINITION | ACCESSION | VERSION | KEYWORDS | SOURCE | ORGANISM | REFERENCE | TITLE | JOURNAL | COMMENT |
|----------|----------|---|-----------|---------|-------------|--------|----------------------|------------------------------------|-------|---------|---------|
| 1 | BU665283 | c1129b12.21 Hembase; Erythroid Precursor Cells (LCB:cl library) | BU665283 | 1 | GI:23377470 | EST | Homo sapiens (human) | 235 bp mRNA linear EST 30-SEP-2002 | | | |
| 5 | 95.2 | 31.7 | 596 | 12 | BG221766 | | | | | | |
| 6 | 94.8 | 31.6 | 422 | 29 | A1833009 | | | | | | |
| 7 | 94.8 | 31.6 | 495 | 28 | AQ236857 | | | | | | |
| 8 | 94.8 | 31.6 | 666 | 28 | B88182 | | | | | | |
| 9 | 94.6 | 31.5 | 570 | 29 | AG019309 | | | | | | |
| 10 | 94.6 | 31.5 | 679 | 29 | AG019484 | | | | | | |
| 11 | 93.8 | 31.3 | 758 | 14 | CA428264 | | | | | | |
| 12 | 93 | 31.0 | 435 | 13 | BQ341376 | | | | | | |
| 13 | 93 | 31.0 | 748 | 10 | BE742023 | | | | | | |
| 14 | 92.6 | 30.9 | 435 | 10 | BE742023 | | | | | | |
| 15 | 92.4 | 30.8 | 276 | 10 | BF916517 | | | | | | |
| 16 | 92.4 | 30.8 | 660 | 28 | AQ021972 | | | | | | |
| 17 | 92.2 | 30.7 | 314 | 9 | AA828637 | | | | | | |
| 18 | 91.8 | 30.6 | 343 | 28 | AQ286677 | | | | | | |
| 19 | 91.4 | 30.5 | 735 | 28 | AQ053384 | | | | | | |
| 20 | 88.8 | 29.6 | 367 | 28 | AQ319539 | | | | | | |
| 21 | 88.8 | 29.6 | 619 | 13 | EX487715 | | | | | | |
| 22 | 88.8 | 29.6 | 673 | 28 | BZ608176 | | | | | | |
| 23 | 87.8 | 29.3 | 607 | 28 | AQ554450 | | | | | | |
| 24 | 87.4 | 29.1 | 319 | 9 | AW088871 | | | | | | |
| 25 | 87.4 | 29.1 | 745 | 13 | BU564170 | | | | | | |
| 26 | 87.4 | 29.1 | 935 | 28 | AQ899972 | | | | | | |
| 27 | 86.4 | 28.8 | 566 | 28 | AQ346062 | | | | | | |
| 28 | 86.4 | 28.8 | 679 | 13 | BU630271 | | | | | | |
| 29 | 86.4 | 28.8 | 952 | 28 | AQ743668 | | | | | | |
| 30 | 86.2 | 28.7 | 453 | 10 | BE160516 | | | | | | |
| 31 | 86.2 | 28.7 | 501 | 9 | AA722505 | | | | | | |
| 32 | 86.2 | 28.7 | 1051 | 10 | BQ026259 | | | | | | |
| 33 | 86 | 28.7 | 452 | 9 | AA668273 | | | | | | |
| 34 | 86 | 28.7 | 845 | 14 | CB993149 | | | | | | |
| 35 | 85.2 | 28.4 | 371 | 9 | A1148153 | | | | | | |
| 36 | 85.2 | 28.4 | 787 | 28 | BZ599533 | | | | | | |
| 37 | 84.6 | 28.2 | 768 | 28 | BZ610312 | | | | | | |
| 38 | 84.4 | 28.1 | 314 | 9 | A1251576 | | | | | | |
| 39 | 84.4 | 28.1 | 872 | 10 | BF680286 | | | | | | |
| 40 | 84.4 | 28.1 | 896 | 13 | BQ924330 | | | | | | |
| 41 | 84.2 | 28.1 | 177 | 10 | AW834029 | | | | | | |
| 42 | 84.2 | 28.1 | 370 | 13 | BX485214 | | | | | | |
| 43 | 84.2 | 28.1 | 630 | 29 | AG159361 | | | | | | |
| 44 | 84.2 | 28.1 | 668 | 29 | AG159359 | | | | | | |
| 45 | 84 | 28.0 | 331 | 9 | AL135357 | | | | | | |

BU665283 235 bp mRNA linear EST 30-SEP-2002
c1129b12.21 Hembase; Erythroid Precursor Cells (LCB:cl library)
BU665283
Homo sapiens cDNA clone c1129b12 5', mRNA sequence.
BU665283.1 GI:23377470
EST
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 235)
Gublin, A.N., Lee, Y.T., Bouffard, G.G. and Miller, J.L.
Gene Expression in Human Erythroid Precursor Cells
Unpublished (2002)
Contact: Jeffrey L. Miller
Laboratory of Chemical Biology
National Institute of Diabetes and Digestive and Kidney Diseases
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
20892, USA
Tel: 301 402 2373
Fax: 301 435 5148
Email: jlmf@nih.gov
The 'cl' library was constructed by Alexander Gublin, Ph.D. in the
Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
analyses by National Institutes of Health Intramural Sequencing

Center (NISC). More information available at:

<http://hembase.niddk.nih.gov>

Plate: 129 row: b column: 12

Seq primer: 5' lambda-TripLex2 Sequencing Primer.

FEATURES

Source

1. 235
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="cl129b12"
/sex="unknown"
/issue_type="blood"
/cell_type="Erythroid Precursor Cells"
/cell_line="Primary Culture of Peripheral Blood Mononuclear Cells"
/dev_stage="Precursor erythroblasts; GFA++"
/lab_host="DH5alpha"
/clone_id="Hembase; Erythroid Precursor Cells (LCB:cl129b12)"
/note="Organ: blood; Vector: pTriplex2; Site 1: SfiI; Site 2: SfiI. A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA library construction kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1 umol/L peptide nucleic acid (PNA) oligos (N-terminal) -biotin-GTC-CAC-CCG-AGG-CTT-G-(C-terminal) and (N-terminal) -biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal). Synthesized cDNA was digested with SfiI and site-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC; <http://www.nisc.nih.gov/>)."

ORIGIN

Query Match 34.7%; Score 104.2; DB 13; Length 235;
Best Local Similarity 75.1%; Pred. No. 1.5e-08;
Matches 130; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 3 GGGCCTGATCCAGCTACTCAGAGGCTGAGGAGAGATCCGGAGCCTGACAG 62
DB 5 GGGCGGATATCCAGCTACTTGAAGGCTGAGGAGAGATCTTGAACCCAGAGA 64
QY 63 TCTGCTGAGCCTGGAGGTTGAGGCTACAGTAAGCAAGTCAATGCAATCTTCA 122
DB 65 ATTGCTTGAACCTGGAGGCGGAGGTTGAGTGAAGCCAAAGATCTTGCAATGCAATCCAG 124
QY 123 CCTGGGCGCAAAAGTGAAGACCGTAAACAAAAAATTTAAAGAAA 175
DB 125 CCTGGGTGACAGATCGAGATCCGCTTAAAAAATTTAAAGAAA 177

RESULT 2
LOCUS AQ0383054 617 bp DNA linear GSS 21-MAY-1999
DEFINITION RPI11-139K16-TV RPI1-11 Homo sapiens genomic clone RPI1-11-139K16, genomic survey sequence.
ACCESSION AQ0383054
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 617)
AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.
TITLE Use of BAC End Sequences from Library RPI1-11 for Sequence-Ready Map Building (1997)
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: RPI11-139K16-TV

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics
The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbe@tigr.org

Clones are derived from the human BAC library RPI1-11. For BAC

library availability, please contact Pieter de Jong

(pieter@jlong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from

Research Genetics (<http://resgen.com>). BAC end search page:

http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html

Seq primer: SP6

Class: BAC ends.

FEATURES

Source

1. 617
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7553247"
/db_xref="taxon:9606"
/clone="RPI1-11-139K16"
/sex="Male"
/cell_type="Lymphocytes"
/clone_id="RPI1-11"
/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; RPI11 Human Male BAC library"

ORIGIN

Query Match 33.9%; Score 101.8; DB 28; Length 617;
Best Local Similarity 69.2%; Pred. No. 2.7e-08;
Matches 139; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 1 GGGGCGCTGATCCAGCTACTCAGAGGCTGAGGAGAGATCCGGAGCCTGGCA 60
DB 490 GGGCACCTGTATTCACAGCTACTGAGGAGGCTGAGGAGAGATCCCTTGAAGCGAGA 431
QY 61 GATCTGCTGAGCCTGGAGGTTGAGGCTACAGTAAGCCAAAGATCATCATCTTTC 120
DB 430 GAATGCTTGAACCCAGAGAGCGAGAGTGAAGTAAGTGAATCAACCACTGCACTCC 371
QY 121 AGCCTGGCGCAAAAGTGAAGCCGTAAACAAAAAATTTAAAGAAATTTAG 180
DB 370 AGCCTGGGTGAAGAGGTGAAGTCCATCTCAAAAAAATTTAAAGAAATTTAG 311
QY 181 ATCAAGATCCAACTGTAAAA 201
DB 310 AAGGAAAAAGAAAGTGAAGAGA 290

RESULT 3
LOCUS AQ035617 509 bp DNA linear GSS 11-JUL-1998
DEFINITION CIT-HSP-2320N18-TR CIT-HSP Homo sapiens genomic clone 2320N18, genomic survey sequence.
ACCESSION AQ035617
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 509)
AUTHORS Adams, M.D., Rounley, S.D., Zhao, S., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and Venter, J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)
JOURNAL Unpublished (1998)
COMMENT Other_GSSs: CIT-HSP-2320N18-TV
The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208

Email: mdamas@cligr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.cligr.org/cdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES

source

Location/Qualifiers
1..509
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2320N18"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

ORIGIN

Query Match 32.5%; Score 97.6; DB 28; Length 509;
Best Local Similarity 70.7%; Pred. No. 1.5e-07;
Matches 130; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 5 GCCTGAATCCAGCTACTCAGAGAGCTGAGGACGAGATCCGGAGCCTGGCAGATC 64
DB 343 GCTTGTGTCACAGTACTCAGAGAGCTGAGGAGGAGATCAGTGAAGCTTGGAGGAT 284
QY 65 TGCCTGAGCCTGGAGGTTGAGGCTACAGTAAGCCAGATCATGCGACTATATCTCAGCC 124
DB 283 CATTTGAACCTCGAGGTTGAGGCTGAGTGAAGCATGATTAACATTAACCTCCAGCC 224
QY 125 TGGCGGACAAAGTGAAGCCGTACAAAATAAATTAAAAAAGAAATTATGATCA 184
DB 223 TGGATGACAGAGAGAGACCTGTCTTAAAAAAGAAAAAATTTAGATGAGAT 164
QY 185 AGAT 188
DB 163 AGTT 160

RESULT 4
A0531557/c 526 bp DNA linear GSS 18-MAY-1999
LOCUS
DEFINITION
RPCI-11-349018.TV RPCI-11 Homo sapiens genomic clone
ACCESSION
A0531557
VERSION
A0531557.1 GI:4839339
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 526)
Zhao, S., Adams, M.D., Nieman, W., Malek, J., de Jong, P. and
Venier, J.C.
TITLE
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
JOURNAL
COMMENT
Unpublished (1997)
Other_GSSs: RPCI-11-349018.TV
Contact: Shaying Zhao, William Nieman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@cligr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@cligr.med.buffalo.edu). Clones may be purchased from
BACpac Resources (http://bacpac.med.buffalo.edu/ordering) or from

Research Genet cs (info@resgen.com). BAC end search page:
http://www.cligr.org/cdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.

FEATURES

source

Location/Qualifiers
1..526
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7633985"
/db_xref="taxon:9606"
/clone="RPCI-11-349018"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"

ORIGIN

Query Match 32.3%; Score 97; DB 28; Length 526;
Best Local Similarity 67.7%; Pred. No. 1.9e-07;
Matches 136; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 1 GGGGGCCTGTAATCCAGCTACTCAGAGAGCTGAGGACGAGATCCGGAGCCTGGCA 60
DB 488 GGGCACCCTGTATCCAGCTACTGAGGAGGCTTGAAGCAGAGATCGTTGAACGAGTA 429
QY 61 GATTCGCTGAGCCTGGAGGTTGAGGCTACAGTAAGCCAGATCATGCGACTATCTTC 120
DB 428 GAATCGCTGAACCCAGAGAGGCGAAGGTTGCGATGATGATCAACCGCTGCACCTCC 369
QY 121 AGCCTGGGCGACAAAGTGAAGCCGTACAAAATAAATAAATTAAAAAAGAAATTAG 180
DB 368 AGCCTGGGCGACAGGTTGAGGCTGAGATCCATCTCAAAAAAATAAAGAAATTAAGA 309
QY 181 ATCAAGATCCACTGTAATAA 201
DB 308 AAGAAAAAGAAAGTGAAGAGA 288

RESULT 5
BG221766/c 596 bp mRNA linear EST 21-APR-2001
LOCUS
DEFINITION
RST41581 Athereys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION
BG221766
VERSION
BG221766.1 GI:13747787
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 596)
Harrington, J.V., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,
Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Bozzer, S.,
Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Colgren, K., Lo, K.,
Offenbacher, J., Danzig, J. and Ducar, M.
TITLE
Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL
COMMENT
Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE
PUBMED
21227151
11329013

CONTACT: Scott J. Cain

Athereys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA

Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athereys.com

High quality sequence stp: 550.

FEATURES

source

Location/Qualifiers
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/cell_line="HT1080"
/clone_lib="Athenrys PAGE Library"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 31.7%; Score 95.2; DB 12; Length 596;
Best Local Similarity 72.1%; Pred. No. 3.7e-07;
Matches 124; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1 GGGGGCTGTATCCAGCTACTCTCAGAGGCTGAGGAGGATCCGGAGGCTGCA 60
DB 286 GGGGGCTGTATCCAGCTACTCTCAGAGGCTGAGGAGGATCCGGAGGCTGCA 227
QY 61 GATCTGCTGAGCTGGAGGCTTGAAGCTACAGTACCAAGATCATCTCCAGTACTTC 120
DB 226 AATAGGCGTGAACCCGGGAGCGGAGCTTGACGTAGCCAGATCGCCCACTGACATCC 167
QY 121 AGCTTGGGCGCAAGAGTGAGACCTTAACAAAAAATTTTAAAAAG 172
DB 166 AGCTTGGGCGCAAGAGTGAGACCTTCAAAAAAAGAG 115

RESULT 6 422 bp mRNA linear EST 21-DEC-1999
A1832009 wj99a04.x1 NCI CGAP Lym12 Homo sapiens CDNA clone IMAGE:2410926 3'
LOCUS similar to contains Alu repetitive element; contains element MER22
DEFINITION repetitive element; mRNA sequence.
ACCESSION A1832009
VERSION A1832009.1 GI:5452680
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 422)
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Life Technologies catalog #: 11547-015
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.lbl.gov/dbp/image/image.html
Insert Length: 538 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 404.

FEATURES

1..422
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2410926"
/issue_type="Lymphoma, follicular mixed small and large
cell"
/lab_host="DH10B"
/clone_lib="NCI CGAP Lym12"
/note="Organ: Lymph node; Vector: pCMV-SportE; Site 1:
SalI; Site 2: NotI; Cloned unidirectionally. Primer:
Oligo dt. Average insert size 1.25 kb. Life Technologies
catalog #: 11547-015"

ORIGIN

Query Match 31.6%; Score 94.8; DB 9; Length 422;
Best Local Similarity 72.4%; Pred. No. 5e-07;

Matches 123; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 5 GCCTGTATCCAGCTACTCTCAGAGGCTGAGGAGGATCCGGAGGCTGGAGATC 64
DB 152 GCTTGTAGTCCAGCTACTCTGAGGCGGAGGAGATGCTTGAAGGAGAGAT 211
QY 65 TGCTTGAAGCTGGAGGCTTGAAGGCTAAGTACCAAGTCAAGCAGTACTTACGAC 124
DB 212 TGCTTGAAGCTGGAGGCTTGAAGGCTAAGTACCAAGTCAAGCAGTACTTACGAC 271
QY 125 TGGGCGCAAGAGTGAACCGTAAACAAAAAATTTAAAAAGAA 174
DB 272 TGGGTACAGAGTGAACCTCTCTCAAAAAAAGAA 321

RESULT 7 495 bp DNA linear GSS 21-APR-1999
A0236857/c
LOCUS RPI11-71K16.TU RPI1-11 Homo sapiens genomic clone RPI1-11-71K16,
DEFINITION genomic survey sequence.
ACCESSION A0236857
VERSION A0236857.1 GI:3669148
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 495)
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P., and Venter, J.C.
TITLE Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPI1-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.bufileo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.bufileo.edu/ordering>) or from
Research Genetics (<http://inforesgen.com>). BAC end search page:
http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

FEATURES

1..495
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7527135"
/db_xref="taxon:9606"
/clone="RPI1-11-71K16"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPI1-11"
/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;
RPI11 Human Male BAC Library"

ORIGIN

Query Match 31.6%; Score 94.8; DB 28; Length 495;
Best Local Similarity 65.7%; Pred. No. 4.7e-07;
Matches 138; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1 GGGGGCTGTATCCAGCTACTCTCAGAGGCTGAGGAGGATCCGGAGGCTGGCA 60
DB 251 GGGGGCTGTATCCAGCTACTCTCAGAGGCTGAGGAGGATCCGGAGGCTGGCA 192
QY 61 GATCTGCTGAGCTGGAGGCTTGAAGCTACAGTACCAAGTCAAGTCAAGTACTTC 120
DB 191 GAATCATTTGAACCCAGAGGAGGAGGTTGACGTGAGCCAAAGTGCCTCATGTATTC 132

QY 121 AGCTTGGGCGCAAAAGTGAACCGGTACAAAAAAATTTTAAAAAGAAATTTAG 180
 |||||
 Db 131 AGCTTGGGCGCAAGTGAAGACTGTGTCTCAAGAAAAAAACCAACAAATTTAT 72
 |||||
 QY 181 ATCAAGATCCAACTGTAAAAAGTGGCCTAA 210
 |||||
 Db 71 TCTGCAATATGAGGCAAGAGTGACTCA 42
 |||||

RESULT 8
 B88182 666 bp DNA linear GSS 09-APR-1999
 LOCUS RPCI11-23114.TPB RPCI-11 Homo sapiens genomic clone RPCI-11-23114,
 DEFINITION genomic survey sequence.
 ACCESSION B88182
 VERSION B88182.1 GI:2929314
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 666)
 Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Baes,S., Linher,K.,
 Golden,K., Berry,K., Granger,D., Sub.E., Wible,C., de Jong,P. and
 Venter,J.C.
 TITLE Use of BAC End Sequences for Sequence-Ready Map Building (1998)
 JOURNAL Unpublished (1998)
 COMMENT Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieterdejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
 Research Genetics (info@resgen.com). BAC end search page:
 http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html
 Seq primer: SP6
 Class: BAC ends.
 Location/Qualifiers
 source 1..666
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="GDB:7508725"
 /db_xref="taxon:9606"
 /clone="RPCI-11-23114"
 /sex="Male"
 /cell_type="Lymphocytes"
 /clone_lib="RPCI-11"
 /note="Vector: PBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
 RPCI11 Human Male BAC Library"

ORIGIN
 Query Match 31.6%; Score 94.8; DB 28; Length 666;
 Best Local Similarity 65.7%; Pred. No. 4.2e-07;
 Matches 18; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1 GGGGGGCTGTAAATCCAGCTACTCAGAGGCTGAGCAGAGATCCGGGAGCCTGGGA 60
 |||||
 Db 210 GGGGTGCTGTAAATCTAGCTACTCAGAGGCTGAGCAGAGATCTATTGACCAAGGA 151
 |||||
 QY 61 GATCTGCTGAGGCTGGAGGTTGAGGCTACAGTAAAGCAAGTATGCGCAGTACTTC 120
 |||||
 Db 150 GAATCTATTGAACCCGAGGAGCGAGAGTTGCACTGAGCCCAAGTTGCCCATGTAATTC 91
 |||||
 QY 121 AGCTTGGGCGCAAAAGTGAACCGGTAAACAAAAAAATTTTAAAAAGAAATTTAG 180
 |||||
 Db 90 AGCTTGGGCGCAAGTGAAGACTGTGTCTCAAGAAAAAAACCAACAAATTTAT 31
 |||||
 QY 181 ATCAAGATCCAACTGTAAAAAGTGGCCTAA 210
 |||||

Db 30 TCTGCAATATGAGGCAAGAGTGACTCA 1
 |||||

RESULT 9
 AG019309 570 bp DNA linear GSS 09-JAN-2003
 LOCUS AG019309
 DEFINITION Homo sapiens genomic DNA, 21q region, clone: B102017 A044(-21),
 genomic survey sequence.
 ACCESSION AG019309
 VERSION AG019309.1 GI:6045253
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 570)
 Hattori,M., Fujiyama,A., Ishii,K., Toyoda,A., Taylor,T.,
 Park,H.-S., Yada,T., Watanabe,H. and Sakaki,Y.
 Homo sapiens genomic DNA, chromosome 21q
 Published Only in Database (1999)
 2 (bases 1 to 570)
 Hattori,M., Fujiyama,A., Ishii,K., Toyoda,A., Taylor,T.,
 Park,H.-S., Yada,T., Watanabe,H. and Sakaki,Y.
 Direct Submission
 Submitted (16-OCT-1999) Masahira Hattori, RIKEN Genomic Sciences
 Center(GSC) c/o Kitasato University; 1-15-1 Kitasato, Sagamihara,
 Kanagawa 228-8555, Japan (E-mail:hattori@gsc.ims.u-tokyo.ac.jp,
 Tel:81-42-778-9923, Fax:81-42-778-9924)
 Location/Qualifiers
 source 1..570
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q"
 /clone="B102017 A044(-21)"

ORIGIN
 Query Match 31.5%; Score 94.6; DB 29; Length 570;
 Best Local Similarity 75.2%; Pred. No. 4.8e-07;
 Matches 118; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 5 GCGCTATATCCAGCTACTCAGAGGCTGAGGAGGAGGATCCGGAGCCTGGAGATC 64
 |||||
 Db 266 GCGCTATATCCAGCTACTCAGAGGCTGAGGAGGAGGATCCCTTGGAGCCCGGATGT 207
 |||||
 QY 65 TGCTGAGGCTGGAGGTTGAGGCTACAGTAAAGCAAGATCATGCAAGTACTTCAGCC 124
 |||||
 Db 206 GCGTTCAAGCTTGGAGTATTAGGCTGAGTGAAGCCGCTATGTCACCTGACTCCAGCC 147
 |||||
 QY 125 TGGGCGCAAAAGTGAACCGGTAAACAAAAAAATTTTAAAAAGAAATTTAG 161
 |||||
 Db 146 TGGGTGCGCAAGTGAAGACTGTGTCTCAAGAAAAAAATTTTAAAAAGAAATTTAG 110
 |||||

RESULT 10
 AG019484 679 bp DNA linear GSS 09-JAN-2003
 LOCUS AG019484
 DEFINITION Homo sapiens genomic DNA, 21q region, clone: B102017_N009_Fw,
 genomic survey sequence.
 ACCESSION AG019484
 VERSION AG019484.1 GI:6045428
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 679)
 Hattori,M., Fujiyama,A., Ishii,K., Toyoda,A., Taylor,T.,
 Park,H.-S., Yada,T., Watanabe,H. and Sakaki,Y.
 Homo sapiens genomic DNA, chromosome 21q
 Published Only in Database (1999)

REFERENCE 2 (bases 1 to 679)
AUTHORS Hattori,M., Fujiyama,A., Ishii,K., Toyoda,A., Taylor,T.,
Park,H.-S., Yada,T., Matanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-1999) Masahira, Hattori, RIKEN Genomic Sciences
Center (GSC), c/o Kitasato University, 1-15-1 Kitasato, Sagamihara,
Kanagawa 228-8555, Japan (E-mail:hattori@gsc.ims.u-tokyo.ac.jp,
Tel:81-42-778-9923, Fax:81-42-778-9924)

FEATURES
source
1. 679
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q"
/clone="B102017_N009_Fw"

ORIGIN

Query Match 31.5%; Score 94.6; DB 29; Length 679;
Best Local Similarity 75.2%; Pred. No. 4.5e-07;
Matches 118; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 5 GCCTGTAATCCAGCTACTCTACAGAGCGCTGAGCGAGAGGATCCGGAGCCTGGCAATC 64
DB 421 GCCTGTAATCCAGCTACTCTACAGAGCGCTGAGCGAGAGGATCCGGAGCCTGGCAATC 362
QY 65 TGCCTGAGCGCTGGAGGTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 124
DB 361 GCCTGAGCGCTGGAGGTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 302
QY 125 TGGCGCAAAAGTGAACCGTACCAAAAAA 161
DB 301 TGGGTGCGAGAGTGAACCGTCTCTCAAAAAA 265

RESULT 11

CA428264/c

LOCUS UI-H-DFO-ber-o-01-0-UI.s1 NCI CGAP_DFO Homo sapiens cDNA clone

DEFINITION UI-H-DFO-ber-o-01-0-UI 3', mRNA sequence.

ACCESSION CA428264

VERSION CA428264.1 GI:24790990

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS 1 (bases 1 to 758)

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov

Tissue Procurement: Dr. Jose Mercuende

cDNA library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, bento-soares@uiowa.edu

The following repetitive elements were found in this cDNA

sequence: 11-240, >ALU (matched complement) 416-674, >ALU (matched

complement)

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers

1. 758

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-H-DFO-ber-o-01-0-UI"

/rname_type="Subchondral Bone"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI CGAP DFO"

/note="Organ: Bone; Vector: pRT73-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

NCI CGAP DFO is a cDNA library containing the following

tissue(s): Subchondral Bone. The library was constructed

according to Bonaldi, Lennon and Soares, Genome Research,

6:791-806, 1996. First strand cDNA synthesis was primed

with an oligo-dT primer containing a Not I site. Double

stranded cDNA was ligated to an EcoR I adaptor, digested

with Not I, and cloned directionally into pRT73-Pac

vector. The oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is

located between the Not I site and the (dT)18 tail. The

sequence tag for this library is GTTAGCGTC.

TAG TISSUE=Subchondral bone

TAG LIB=UI-H-DFO

TAG_SEQ=GTTAGCGTC"

ORIGIN

Query Match 31.3%; Score 93.8; DB 14; Length 758;
Best Local Similarity 75.7%; Pred. No. 5.9e-07;
Matches 131; Conservative 0; Mismatches 37; Indels 5; Gaps 1;

QY 1 GGGGCGCTGTAATCCAGCTACTCTACAGAGCGCTGAGCGAGAGGATCCGGAGCCTG--- 57
DB 173 GGGTGCCTGTAATCCAGCTACTCTACAGAGCGCTGAGCGAGAGGATCCGGAGCCTGAGG 114
QY 58 --GCAATCTGCTGAGCTTGGAGGTTGAGGCTACAGTAAGCCCAAGATCAGCCAGTAT 115
DB 113 CAGGAGATCACTTGAACCTAGAGGCAAGAGGTTCAAGAGCCCAAGATCAGCCAGTATTC 54
QY 116 ACTTCAGCGCTGGGCGCAAAAGTGAACCGTACCAAAAAAATTAAAA 168
DB 53 ACTTCAGCGCTGGGCGCAAAAGTGAACCGTACCTCTCAAAAAA 1

RESULT 12

BQ341376

LOCUS BQ341376 361 bp mRNA linear EST 20-MAY-2002

DEFINITION R05-NN1065-300600-022-C04 NN1065 Homo sapiens cDNA, mRNA sequence.

ACCESSION BQ341376

VERSION BQ341376.1 GI:21003349

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 361)

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Brites,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,G.S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE

COMMENT

PUBMED

CONTACT: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the PABSP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

<http://www.ludwig.org.br/scripts/gethtml2.pl?l=R05&t2=R05-NN1065-300600-022-C04&t3=2000-06-30&t4=1>

Seq primer: puc 16 forward

FEATURES

source

High quality sequence stop: 381.

Location/Qualifiers

1..381

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_idb="NN1065"

/note="Organ: nervous normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent Application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match

Best Local Similarity 68.3%; Pred. No. 1.1e-06; Matches 129; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 5 GCCTTATCCAGTCTACAGAGGCTGAGGAGATCCGCGAGCTTGGCAATC 64
DB 62 GCTTGAGCTCTAGTCTGAGGAGCTGAGCTGAGCTTGAAGCTTGGGAGTT 121
QY 65 TGCCGAGCCTGGAGGTTGAGGCTACAGTAGCCAGATCATGCGATATCTTACGCC 124
DB 122 TCTTGAGCCTGGAGGCTGAGGCTCCAGTAGCCATATTTGTCATCTGCATTCAGCC 181
QY 125 TGGGCGACAAAGTGAGCCGTACACAAAAAATTTAAAAAAGAAATTTAGATCA 184
DB 182 TGGGTAACTGAGCGAGACTGTCTCTCAACAAATTAATAAAGCTGAAGAATTCATCA 241
QY 185 AGATCCAAAC 193
DB 242 CCAACAAAC 250

RESULT 13

AA515045/c

LOCUS

AA515045 435 bp mRNA linear EST 19-AUG-1997

DEFINITION n667h05.s1 NCI CGAP L1P2 Homo sapiens cDNA clone IMAGE:39897

similar to contains Alu repetitive element; contains element TARI

Repetitive element; mRNA sequence.

ACCESSION AA515045

VERSION AA515045.1 GI:2254645

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 435)

NCI-CCAP http://www.nci.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: David B. Krizman, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

clone distribution: NCI-CCAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.livni.gov/bbtp/image/image.html

Insert Length: 947 Std Error: 0.00

Seq primer: -40m3 fwd. ET from Amersham

High quality sequence stop: 375.

Location/Qualifiers

1..435

/organism="Homo sapiens"

/mol_type="mRNA"

FEATURES

source

/db_xref="taxon:9606"
/clone_idb="IMAGE:39897"
/tissue_type="liposarcoma"
/lab_host="DH10B"
/clone_idb="NCI CGAP L1P2"
/note="Vector: pAMP10; mRNA made from liposarcoma, cDNA made by oligo-dt priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."

ORIGIN

Query Match

Best Local Similarity 68.3%; Pred. No. 1e-06; Matches 129; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 5 GCCTTATCCAGTCTACAGAGGCTGAGGAGATCCGCGAGCTTGGCAATC 64
DB 425 GCTTGAGCTCTAGTCTGAGGAGCTGAGCTGAGGAGTTCACTTGAAGCTTGGGAGTT 366
QY 65 TGCCGAGCCTGGAGGTTGAGGCTACAGTAGCCAGATCATGCGATATCTTACGCC 124
DB 365 TCTTGAGCCTGGAGGCTGAGGCTCCAGTAGCCATATTTGTCATCTGCATTCAGCC 306
QY 125 TGGGCGACAAAGTGAGCCGTACACAAAAAATTTAAAAAAGAAATTTAGATCA 184
DB 305 TGGGTAACTGAGCGAGACTGTCTCTCAACAAATTAATAAAGCTGAAGAATTCATCA 246
QY 185 AGATCCAAAC 193
DB 245 CCAACAAAC 237

RESULT 14

BE742023/c

LOCUS

BE742023 748 bp mRNA linear EST 15-SEP-2000

DEFINITION 601594512F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:394838 5',

mRNA sequence.

ACCESSION BE742023

VERSION BE742023.1 GI:10156015

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 748)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.livni.gov

plate: L10M811 row: h column: 03

High quality sequence start: 29

High quality sequence stop: 745.

Location/Qualifiers

1..748

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone_idb="IMAGE:394838"

/tissue_type="adipocarcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/clone_idb="NIH MGC 9"

/note="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dt priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGACGAG(G). Size-selected >500bp for average

Source

ORIGIN

Db 197 TTCAAGATTATTTTACTGCTGGAATCTTTGTAGAAACCATGAAGAAGCTTTCAAA 256

Qy 78 GCCCTGGGAAGGGCTGCATGAAATTCACTTCCCTCTCTCCAGACAGCTTCGGCCATCCTT 137
Dp 257 GCCCTGGGAAGGGCTGCATGAAATTCACTTCCCTCTCTCCAGACAGCTTCGGCCATCCTT 316

| | | | |
|----|-----|-----|-----|
| QY | 138 | TTG | 140 |
| Db | 317 | TTG | 319 |

6133

| | | | | | |
|------------|---|--------|------|--------|-----------------|
| LOCUS | BG686133 | 743 bp | mRNA | linear | EST 01-MAY-2000 |
| DEFINITION | 602638412F1 NIH_MGC_48 Homo sapiens cDNA IMAGE:4766234 5', mRNA sequence. | | | | |

```

VERSION      EG686133.1  GI:13917530
KEYWORDS
SOURCE       Homo sapiens (human)

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 743)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL Unpublished (1998)
COMMENT Contact: Robert Strausberg, Ph.D.

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CMT626 row: g column: 03
High quality sequence step: 740.

Source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="raxon:9606"
/clone="IMAGE:4766234"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_id="NH_MGC_48"

```

| | | | | |
|---------------------------|-------|--------------------|-----------|-------------|
| Query Match | 39.3% | Score 118.2; | DB 12; | Length 743; |
| Best Local Similarity | 97.6% | Pred. No. 1.2e-11; | | |
| Matches 120; Conservative | 0; | Mismatches 3; | Indels 0; | Gaps 0; |

Oy 18 TTGAGATTATTTTACTGCTGGAACTTTGTAGAAAACCGAAGAATACTTTCAA 77
 ||| |||||
Db 498 TTCGAAGATTATTTTACTGCTGGAATCTTTGTAGAAAACCATGAAGAACCTTTCAA 555

2y 78 GCGTGGGAAGGGCGTCGATGAAAAATTCAGTTGCTCTCTCCAGACAGAGTTCCGGGCATCCTT 137
 Db 558 GCGTGGGAAGGGCGTCGATGAAAAATTCAGTTGCTCTCTCCAGACAGAGTTCCGGGCATCCTT 617

| | | | |
|----|-----|-----|-----|
| QY | 138 | TTG | 140 |
| | | | |
| Db | 618 | TTG | 620 |

5757089

| | | | | |
|------------|---|------|--------|-----------------|
| LOCUS | 820 bp | mRNA | linear | EST 15-MAY-2000 |
| DEFINITION | 602715124P1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4855517 5', mRNA sequence. | | | |

| | | |
|----------|----------------------|-------------|
| VERSION | EG57089.1 | GI:14067742 |
| KEYWORDS | EST. | |
| SOURCE | Homo sapiens (human) | |
| ORGANISM | | |

REFERENCE
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 820)
AUTHORS NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.

JOURNAL COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

BOU

Plate: L1CM1704 row: 0 column: 06
High quality sequence stop: 675.

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="14855517"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/notes="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI. cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GCACGAG (g). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

ORIGIN

Query Match 39.3%; Score 118.2; DB 12; Length 820;
 Best Local Similarity 97.6%; Pred. No. 1.1e-11;
 Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 TTGAAGATTATTTTACTGCTGGAATCTTTGTAGAAAACGAGAAAGACTTTCAA 77
 |||
 DB 505 TTCAAAGATTATTTTACTGCTGGAATCTTTGTAGAAAACGAGAAAGACTTTCAA 564
 |||
 QY 78 GCGTGGGAAGGGCGCTGATGAAATTCAGTTCCTCTCCAGACAGCTTCGGCGATCCTT 137
 |||
 DB 565 GCGTGGGAAGGGCGCTGATGAAATTCAGTTCCTCTCCAGACAGCTTCGGCGATCCTT 624
 |||
 QY 138 TTG 140
 |||
 DB 625 TTG 627

RESULT 4

BG758510 872 bp mRNA linear EST 15-MAY-2001
 LOCUS 602712721 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853069 5',

DEFINITION

mRNA sequence.

ACCESSION

BG758510

VERSION

BG758510.1 GI:14069163

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens (human)

REFERENCE

1 (bases 1 to 872)

AUTHORS

NIH-MGC http://mgs.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov

FEATURES

CDNA Library Preparation: Ling Hong/Rubin Laboratory

source

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

FEATURES

DNA Sequencing by: Inocyte Genomics, Inc.

source

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov

FEATURES

Plate: LNCM1698 row: 1 column: 06

source

High quality sequence step: 836.

FEATURES

Location/Qualifiers

source

1..872

FEATURES

/organism="Homo sapiens"

source

/mol_type="mRNA"

FEATURES

/db_xref="taxon:9606"

source

/clone="IMAGE:4853069"

FEATURES

/tissue="Primary B-cells from tonsils (cell line)"

source

/lab_host="DH10B (phage-resistant)"

FEATURES

/clone_1ib="NIH_MGC_48"

source

/note="Organ: B-cells; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 39.3%; Score 118.2; DB 12; Length 872;
 Best Local Similarity 97.6%; Pred. No. 1.1e-11;
 Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 TTGAAGATTATTTTACTGCTGGAATCTTTGTAGAAAACGAGAAAGACTTTCAA 77
 |||
 DB 486 TTCAAAGATTATTTTACTGCTGGAATCTTTGTAGAAAACGAGAAAGACTTTCAA 545
 |||

QY

78 GCGTGGGAAGGGCGCTGATGAAATTCAGTTCCTCTCCAGACAGCTTCGGCGATCCTT 137

DB

546 GCGTGGGAAGGGCGCTGATGAAATTCAGTTCCTCTCCAGACAGCTTCGGCGATCCTT 605

QY

138 TTG 140

DB

606 TTG 608

RESULT 5

BQ065440 953 bp mRNA linear EST 02-APR-2002

LOCUS

AGENCOURT 6855061 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5929977

DEFINITION

5', mRNA sequence.

ACCESSION

BQ065440

VERSION

BQ065440.1 GI:19894486

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens (human)

REFERENCE

1 (bases 1 to 953)

AUTHORS

NIH-MGC http://mgs.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov

FEATURES

CDNA Library Preparation: Rubin Laboratory

source

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

FEATURES

DNA Sequencing by: Agencourt Bioscience Corporation

source

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov

FEATURES

Plate: LNCM2108 row: P column: 10

source

High quality sequence step: 634.

FEATURES

Location/Qualifiers

source

1..953

FEATURES

/organism="Homo sapiens"

source

/mol_type="mRNA"

FEATURES

/db_xref="taxon:9606"

source

/clone="IMAGE:5929977"

FEATURES

/tissue="Lymphoma, cell line"

source

/lab_host="DH10B (phage-resistant)"

FEATURES

/clone_1ib="NIH_MGC_99"

source

/note="Organ: lymph; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 39.3%; Score 118.2; DB 13; Length 953;
 Best Local Similarity 97.6%; Pred. No. 1e-11;
 Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 TTGAAGATTATTTTACTGCTGGAATCTTTGTAGAAAACGAGAAAGACTTTCAA 77
 |||
 DB 484 TTCAAAGATTATTTTACTGCTGGAATCTTTGTAGAAAACGAGAAAGACTTTCAA 543
 |||
 QY 78 GCGTGGGAAGGGCGCTGATGAAATTCAGTTCCTCTCCAGACAGCTTCGGCGATCCTT 137
 |||
 DB 544 GCGTGGGAAGGGCGCTGATGAAATTCAGTTCCTCTCCAGACAGCTTCGGCGATCCTT 603
 |||
 QY 138 TTG 140
 |||
 DB 604 TTG 606

RESULT 6
LOCUS BQ055935 1052 bp mRNA linear EST 29-MAR-2002
DEFINITION AGENCOURT 6796291 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5808181
5', mRNA sequence.
ACCESSION BQ055935
VERSION BQ055935.1 GI:19815262
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1052)
NIH-MGC <http://mgs.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
AUTHORS Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LICM2051 Row: m Column: 14
High quality sequence stop: 665.
Location/Qualifiers
1. 1052
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5808181"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
library."

ORIGIN
Query Match 39.3%; Score 118.2; DB 13; Length 1052;
Best Local Similarity 97.6%; Pred. No. 9.8e-12;
Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
DB 18 TTTGAAGATTATTTACTGCTGGAATCTTTGTAGAAACCAAGAAAGACTTTCAA 77
|||
484 TTCAAGATTATTTACTGCTGGAATCTTTGTAGAAACCAAGAAAGACTTTCAA 543
|||
QY 78 GCTGGGAAGGCTGCATGAAATTCAGTTCCTCTCCAGACAGCTTCGGGCACTCTT 137
|||
DB 544 GCCTGGGAAGGCTGCATGAAATTCAGTTCCTCTCCAGACAGCTTCGGGCACTCTT 603
|||
QY 138 TTG 140
|||
DB 604 TTG 606

RESULT 7
LOCUS AL559877 1201 bp mRNA linear EST 31-MAY-2003
DEFINITION AL559877 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
clone CS0D003YB14 5-PRIME, mRNA sequence.
ACCESSION AL559877
VERSION AL559877.2 GI:31284008
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1201)
NIH-MGC <http://mgs.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (2001)
AUTHORS Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

FEATURES
source
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D003YB14"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_1ib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 39.3%; Score 118.2; DB 9; Length 1201;
Best Local Similarity 97.6%; Pred. No. 9.2e-12;
Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
DB 512 TTCAAGATTATTTACTGCTGGAATCTTTGTAGAAACCAAGAAAGACTTTCAA 571
|||
QY 78 GCTGGGAAGGCTGCATGAAATTCAGTTCCTCTCCAGACAGCTTCGGGCACTCTT 137
|||
DB 572 GCCTGGGAAGGCTGCATGAAATTCAGTTCCTCTCCAGACAGCTTCGGGCACTCTT 631
|||
QY 138 TTG 140
|||
DB 632 TTG 634

RESULT 8
LOCUS BP975166 942 bp mRNA linear EST 22-JUN-2001
DEFINITION 602244657F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4335639 5',
mRNA sequence.
ACCESSION BP975166
VERSION BP975166.1 GI:12342381
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 942)
NIH-MGC <http://mgs.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
AUTHORS Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLCMI207 row: a column: 16
High quality sequence stop: 707.
Location/Qualifiers

FEATURES

1..942
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4335639"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 39.1%; Score 117.6; DB 10; Length 942;
Best Local Similarity 96.8%; Pred. No. 1.3e-11;
Matches 120; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 TTGAAGATTATTTTACTGCTGGAATCTTTGTAGAAACACGAAAGACTTTCAA 77
DB 499 TTCAAAGATTATTTTACTGCTGGAATCTTTGTAGAAACACGAAAGACTTTCAA 558
QY 78 GCTGGGAGGCGTGATGAAATTCAGTTCTCTCCAGACAGCTTGGCGCATCCTT 137
DB 559 GCTGGGAGGCGTGATGAAATTCAGTTCTCTCCAGACAGCTTGGCGCATCCTT 618
QY 138 TTGG 141
DB 619 TGGG 622

RESULT 9
BG757392 693 bp mRNA linear EST 15-MAY-2001
LOCUS 602711022F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4851580 5',
DEFINITION mRNA sequence.
ACCESSION BG757392
VERSION BG757392.1 GI:14068045
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 693)
NIH-MGC <http://img.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLCMI694 row: K column: 05
High quality sequence stop: 693.
Location/Qualifiers

FEATURES

source

1..693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4851580"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 38.7%; Score 116.6; DB 12; Length 693;
Best Local Similarity 96.7%; Pred. No. 2.3e-11;
Matches 119; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 TTGAAGATTATTTTACTGCTGGAATCTTTGTAGAAACACGAAAGACTTTCAA 77
DB 502 TTCAAAGATTATTTTACTGCTGGAATCTTTGTAGAAACACGAAAGACTTTCAA 561
QY 78 GCTGGGAGGCGTGATGAAATTCAGTTCTCTCCAGACAGCTTGGCGCATCCTT 137
DB 562 GCTGGGAGGCGTGATGAAATTCAGTTCTCTCCAGACAGCTTGGCGCATCCTT 621
QY 138 TTG 140
DB 622 TTG 624

RESULT 10
BX464579 853 bp mRNA linear EST 22-MAY-2003
LOCUS BX464579 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
DEFINITION clone CS0DG003YB14 5-PRIME, mRNA sequence.
ACCESSION BX464579
VERSION BX464579.1 GI:31031641
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 853)
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6672.x For
more information about this cluster, see
<http://www.genoscope.cns.fr/>
cgi-bin/cluster.cgi?seq=CS1D5001ZP10QPI&cluster=6672.x. Contact :
Feng Liang Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS1D5001ZP10QPI.
Location/Qualifiers

FEATURES

1..853
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DG003YB14"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"

/note="Vector: PCWVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the PCWVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 38.7%; Score 116.6; DB 13; Length 853;
Best Local Similarity 96.7%; Pred. No. 2.1e-11;
Matches 119; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 TTGAAGATTATTTTACTGCTGGAATCTTTGTAGAAAACGAGAAAGACTTTCAA 77
DB 506 TTCAAGATTATTTTCTGCTGGAATCTTTGTAGAAAACGAGAAAGACTTTCAA 565
QY 78 GCGTGGGAGGCGCTGATGAAATTCAGTTCCTCTCCAGACAGCTTCGGCGATCCTT 137
DB 566 GCGTGGGAGGCGCTGATGAAATTCAGTTCCTCTCCAGACAGCTTCGGCGATCCTT 625
QY 138 TTG 140
DB 626 TTG 628

RESULT 11
BG755005 1034 bp mRNA linear EST 15-MAY-2001
LOCUS 6027115.1F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4851819 5',
DEFINITION mRNA sequence.
ACCESSION BG755005
VERSION BG755005.1 GI:14065658
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1034)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (HLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/HLNL at:
<http://image.llnl.gov>
Plate: LNCM1695 row: e column: 04
High quality sequence stop: 622.

FEATURES

Location/Qualifiers
1..1034
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4851819"
/issue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 38.1%; Score 114.8; DB 12; Length 1034;
Best Local Similarity 91.0%; Pred. No. 3.9e-11;

Matches 122; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 7 TTTTCTTTTGTGAGATTTATTTTACTGCTGGAATCTTTGTAGAAAACGAGAA 66
DB 27 TTTTCTTTTGTGAGATTTATTTTACTGCGGGAATCTTTGTAGAAAACGAGAA 86
QY 67 GAACCTTCAAGCGCGGAGGCGCTGATGAAATTCAGTTCCTCTCCAGACAGCTTC 126
DB 87 GAACCTTCAAGCGCGGAGGCGCTGATGAAATTCAGTTCCTCTCCAGACAGCTTC 146
QY 127 GCGCATCTCTTTG 140
DB 147 GCGCATCTCTTTG 160

RESULT 12
BH302559/c 442 bp DNA linear GSS 30-NOV-2001
LOCUS CH230-100C8.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
DEFINITION CH230-100C8, genomic survey sequence.
ACCESSION BH302559
VERSION BH302559.1 GI:17214967
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 442)
AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsagaye, G., Geer, K., Shwartsbeyn, A., Gebregiorgis, E., Overton, L., Russell, D., Chen, D., Riggs, F., de Jong, P. and Fraser, C.M.
TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment Unpublished (1999)
JOURNAL Other GSSs: CH230-100C8.TV
COMMENT Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230 (<http://www.chori.org/bacpac/rat230.htm>). For BAC library availability, please contact Pieter de Jong (pdj@omni.net). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/or_exting_information.htm). BAC end page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
Plate: 100 row: C column: 8
Seq primer: T7
Class: BAC ends.

FEATURES

Location/Qualifiers
1..442
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-100C8"
/sex="Female"
/cell_type="Brain"
/clone_id="CHORI-230 Segment 1"
/note="Vector: pTABAC2.1; Site_1: EcoRI; Site_2: EcoRI; CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by Pieter de Jong"

ORIGIN

Query Match 38.1%; Score 114.6; DB 28; Length 442;
Best Local Similarity 86.9%; Pred. No. 6.4e-11;
Matches 126; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 TTTTCCTTTTCTTTTGTGAGATTTATTTTACTGCTGGAATCTTTGTAGAAAAC 60
DB 382 TTATGCAATTAATTTCTCTTAAGACTATTTTACTGCGGAATCAATTTGTAGAAAATC 323

QY 61 AGCAAGAACTTTTAAAGCTGGGAGGCTGATGAAATTCAGTTCCTCCAGAC 120
 DB 322 ATGAAAGAACTTTTAAAGCTGGGAGGCTGATGAAATTCAGTTCCTCCAGAC 263
 QY 121 AGCTTGGCGGATCTCTTTGTAG 145
 DB 262 AGCTTGGCGGATCTCTTTGTAG 238

RESULT 13

LOCUS BX402063 1201 bp mRNA linear EST 13-MAY-2003
 DEFINITION BX402063 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 Homo sapiens cDNA clone CS0D012YD18 5-PRIME, mRNA sequence.
 ACCESSION BX402063
 VERSION BX402063.1 GI:30626645
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1201)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqreq@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Email: seqreq@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 6672.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0D012DB09Q1&cluster=6672.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0D012DB09Q1.

FEATURES

source

1..1201
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 /clone="CS0D012YD18"
 /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /cell_line="RAMOS CELL LINE"
 /clone_1b="Homo sapiens B CELLS (RAMOS CELL LINE) COT
 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 35.6%; Score 107.2; DB 13; Length 1201;
 Best Local Similarity 96.8%; Pred. No. 8e-10; 3; Indels 1; Gaps 1;
 Matches 120; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 18 TTGAAGATTATTTTACTGCTGGAATACT-TTGTAGAAAAACGAGAAAGACTTCAA 76
 DB 572 TTGAAGATTATTTTACTGCTGGAATACTGTTGTAGAAAAACGAGAAAGACTTCAA 631
 QY 77 AGCCTGGGAGGCTGATGAAATTCAGTTCCTCTCCAGACAGCTTCGGGCACTCT 136
 DB 632 AGCCTGGGAGGCTGATGAAATTCAGTTCCTCTCCAGACAGCTTCGGGCACTCT 691
 QY 137 TTGG 140
 DB 692 TTGG 695

RESULT 14

BG758815

LOCUS BG758815 843 bp mRNA linear EST 15-MAY-2001
 DEFINITION 602713177P1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853467 5',
 mRNA sequence.
 ACCESSION BG758815
 VERSION BG758815.1 GI:14069468
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 843)
 NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contract: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNCM1699 row: 1 column: 20
 High quality sequence stop: 615.
 Location/Qualifiers
 1..843
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4853467"
 /tissue_type="Primary B-cells from tonsils (cell line)"
 /lab_note="DH10B (phage-resistant)"
 /clone_1b="NIH_MGC_48"
 /note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

FEATURES

source

1..843
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4853467"
 /tissue_type="Primary B-cells from tonsils (cell line)"
 /lab_note="DH10B (phage-resistant)"
 /clone_1b="NIH_MGC_48"
 /note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 35.0%; Score 105.4; DB 12; Length 843;
 Best Local Similarity 91.1%; Pred. No. 2e-09;
 Matches 123; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
 QY 7 TTTTATTTTATTTTGAAGATTATTTTACTGCTGGAATACTTTGTAGAAAAACGAGAA 66
 DB 27 TTTTATTTTATTTTGAAGATTATTTTACTGCTGGAATACTTTGTAGAAAAACGAGAA 86
 QY 67 GAACCTT-CAAGCTGGGAGGCTGATGAAATTCAGTTCCTCTCCAGACAGCTT 125
 DB 87 GAACCTTCAAGGCTGGGAGGCTGATGAAATTCAGTTCCTCTCCAGACAGCTT 146
 QY 126 CGGCGCATCTTTTG 140
 DB 147 CGGCGCATCTTTTG 161

RESULT 15

LOCUS BG144705 522 bp mRNA linear EST 01-FEB-2001
 DEFINITION uc73f07.y1 Soares_mouse_muscle_bcell Mus musculus cDNA clone
 IMAGE:333637 5' similar to TR:Q9NWO Q9NWO ACTIVATION-INDUCED
 CYTIDINE DEAMINASE, mRNA sequence.
 ACCESSION BG144705
 VERSION BG144705.1 GI:12648105
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 522)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

COMMENT

Email: cgapbs@mail.nih.gov

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.

MGI:1077801

Seq primer: -40RP from Gibco

High quality sequence stop: 487.

FEATURES

Location/Qualifiers

1..522

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:333637"

/lab_host="DH10B (phage-resistant)"

/clone_lib="Soares_mouse_MG8_Dcell1"

/note="Organ: Germinal B-cell; Vector: p773D-Pac

(Pharmacia) with a modified polylinker; Site 1: Not I;

Site 2: Eco RI; 1st strand cDNA was primed with a Not I -

oligo(dT) primer [5'

TGTTACCAATCTGAGCTGAGGAGCGCCGCTGTTTCTTTTCTTTTCTTTT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified p773 vector.

Library is normalized; constructed by Bento Soares and

M.Fatima Bonaldi."

ORIGIN

Query Match

33.4%; Score 100.6; DB 12; Length 522;

Best Local Similarity 88.6%; Pred.No.1.8e-08;

Matches 109; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

18 TTGAGATTATTTTCTGCTGGAATCTTTTGTAGAAAACGAGAAAGACTTCAAA 77

338 TTCAAGACTATTTTCTGCTGGAATCTTTGTAGAAAATCGTGAAGAACTTCAAA 397

78 GCCCGGGAAGGCGCGCATGGAATTCAGTCGCTCTCCAGACAGCTCGGCGCATCCTT 137

398 GCCCGGGAAGGCGCGCATGGAATTCAGTCGCTCTCCAGACAGCTCGGCGCATCCTT 457

138 TTG 140

458 TTG 460

Search completed: March 13, 2004, 00:38:56

Job time : 962.585 secs

QY 121 ACCTCGGCGCATCCTTTGGTAAGGGGCTTCCTGCTTTTAAATTTCTTTCTC 180
DB 3252 ACCTCGGCGCATCCTTTGGTAAGGGGCTTCCTGCTTTTAAATTTCTTTCTC 3311
QY 181 TACAGCTTTTGGAGTTTGGTATTTCTTATATTTCTTATTTGTTCATCATCTCA 240
DB 3312 TACAGCTTTTGGAGTTTGGTATTTCTTATATTTCTTATTTGTTCATCATCTCA 3371
QY 241 GTTTCATCTGATGAAGAACTTATTTCTCTCCATCAGCTTTTCTTCTGCTGTTCA 300
DB 3372 GTTTCATCTGATGAAGAACTTATTTCTCTCCATCAGCTTTTCTTCTGCTGTTCA 3431
QY 301 C 301
DB 3432 C 3432

RESULT 2

US-09-966-880A-35
Sequence 35, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Muramatsu, Masamichi
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966,880A
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 35
LENGTH: 11204
TYPE: DNA
ORGANISM: Homo sapiens
US-09-966-880A-35

Query Match 100.0%; Score 301; DB 9; Length 11204;
Best Local Similarity 100.0%; Pred. No. 1e-61;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTCTTTTCTTTTCTTTTGAAGATATTTTACTGCTGAATCTTTGTGAGAAAC 60
DB 8347 TTTTCTTTTCTTTTCTTTTGAAGATATTTTACTGCTGAATCTTTGTGAGAAAC 8406
QY 61 ACGAAAGACTTCAAGGCTGGAAGGCTGCATGAATAATCACTGCTCTCCAGAC 120
DB 8407 ACGAAAGACTTCAAGGCTGGAAGGCTGCATGAATAATCACTGCTCTCCAGAC 8466
QY 121 AGCTCGGCGCATCCTTTGGTAAGGGGCTTCCTGCTTTTAAATTTCTTTCTC 180
DB 8467 AGCTCGGCGCATCCTTTGGTAAGGGGCTTCCTGCTTTTAAATTTCTTTCTC 8526
QY 181 TACAGCTTTTGGAGTTTGGTATTTCTTATATTTCTTATTTGTTCATCATCTCA 240
DB 8527 TACAGCTTTTGGAGTTTGGTATTTCTTATATTTCTTATTTGTTCATCATCTCA 8586
QY 241 GTTTCATCTGATGAAGAACTTATTTCTCTCCATCAGCTTTTCTTCTGCTGTTCA 300
DB 8587 GTTTCATCTGATGAAGAACTTATTTCTCTCCATCAGCTTTTCTTCTGCTGTTCA 8646
QY 301 C 301
DB 8647 C 8647

RESULT 3

US-09-966-880A-7
Sequence 7, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Muramatsu, Masamichi
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966,880A
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 2818
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (80)...(673)
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)...(79)
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (677)...(2818)
US-09-966-880A-7

Query Match 39.3%; Score 118.2; DB 9; Length 2818;
Best Local Similarity 97.6%; Pred. No. 2.5e-18;
Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 TTTGAAGATTTATTTTACTGCTGATTAATCTTTGTGAGAAACGAGAAACTTTCAA 77
DB 500 TTTGAAGATTTATTTTACTGCTGATTAATCTTTGTGAGAAACGAGAAACTTTCAA 559
QY 78 GCCTGGAGAGGCTGCATGAATAATCACTGCTCTCCAGACAGCTTTGGCGCATCCT 137
DB 560 GCCTGGAGAGGCTGCATGAATAATCACTGCTCTCCAGACAGCTTTGGCGCATCCT 619
QY 138 TTG 140
DB 620 TTG 622
RESULT 4
US-09-966-880A-14
Sequence 14, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Muramatsu, Masamichi
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966,880A
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29

NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 116
TYPE: DNA
ORGANISM: Homo sapiens
US-09-966-880A-14

Query Match
Best Local Similarity 100.0%; Pred. No. 2.5e-18;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 ATTATTTTACTGCTGGAATCTTTGTAGAAACCAAGAACTTTCAAGCTGG 84
DB 1 ATTATTTTACTGCTGGAATCTTTGTAGAAACCAAGAACTTTCAAGCTGG 60
QY 85 AAGGGCTGATGAAATTCAGTTCTCTCCAGACGCTTGGGGCATCCTTTTG 140
DB 61 AAGGGCTGATGAAATTCAGTTCTCTCCAGACGCTTGGGGCATCCTTTTG 116

RESULT 5

US-09-966-880A-1
Sequence 1, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Muramatsu, Masamichi
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966,880A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2440
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (93)...(686)
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)...(92)
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (690)...(2440)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(2440)
OTHER INFORMATION: n = A,T,C or G
US-09-966-880A-1

Query Match
Best Local Similarity 33.4%; Score 100.6; DB 9; Length 2440;
Matches 109; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 18 TTGAAGATTATTTTACTGCTGGAATCTTTGTAGAAACCAAGAACTTTCAA 77
DB 513 TTCAAGACTATTTTACTGCTGGAATCTTTGTAGAAACCAAGAACTTTCAA 572
QY 78 GCTGGGAAGGGCTGATGAAATTCAGTTCTCTCCAGACGCTTGGGGCATCCTT 137
DB 573 GCTGGGAAGGGCTGATGAAATTCAGTTCTCTCCAGACGCTTGGGGCATCCTT 632

QY 138 TTG 140
DB 633 TTG 635

RESULT 6

US-09-918-995-31138/C
Sequence 31138, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 31138
LENGTH: 475
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(475)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-31138

Query Match
Best Local Similarity 15.9%; Score 47.8; DB 10; Length 475;
Matches 97; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 83 GGAAGGCTGCATGAAATTCAGTTCTCTCCAGACGCTTGGGGCATCCTTTTGT 142
DB 388 GGGGGGCTTTAAAAAACCAGGGGCTTTCCGGGGGCTTTTCCCGCCCTTTT 329
QY 143 AAGGGCTGCTGCTGCTTTAAATTTCTCTCTCTCTCTCTCTCTCTCTCT 202
DB 328 TTGGGGGCCCCNNNNNTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 269
QY 203 TATATTTCTATATTTCTATATTTCTATATTTCTATATTTCTATATTTCT 262
DB 268 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 209
QY 263 ATT 265
DB 208 TTT 206

RESULT 7

US-09-918-995-10105/C
Sequence 10105, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10105
LENGTH: 506
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(506)

OTHER INFORMATION: n = A,T,C or G
US-09-918-995-10105

Query Match 15.9%; Score 47.8; DB 10; Length 506;
Best Local Similarity 58.6%; Pred. No. 0.084;
Matches 82; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 127 GGGCGATCCTTTGGTAAGGGGCTCCGCGCTTTAAATTTCTTCTTCTACAGT 186
DB 324 GGGCGATCCTTTGGTAAGGGGCTTTTCTTTCTTTCTTTCTTTCTTTCTTT 265
QY 187 CTTTGTGAGTTTGGTAATTTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 246
DB 264 TTTTGTGAGTTTGGTAATTTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 205
QY 247 ATCTGATGAAAAGTTTATTT 266
DB 204 GTTCCATGAAAAGTTTATTT 185

RESULT 8

US-10-424-599-3232/c
Sequence 3232, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 3232
LENGTH: 344
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_102923C.1
US-10-424-599-3232

Query Match 15.8%; Score 47.6; DB 12; Length 344;
Best Local Similarity 54.6%; Pred. No. 0.081;
Matches 95; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 127 GGGCGATCCTTTGGTAAGGGGCTCCGCGCTTTAAATTTCTTCTTCTACAGT 186
DB 198 GGGCGGCTCCAGGAGGGGGGCGCCCTTTTCTTTCTTTCTTTCTTTCTTTCT 139
QY 187 CTTTGTGAGTTTGGTAATTTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 246
DB 138 TTTTGTGAGTTTGGTAATTTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 79
QY 247 ATCTGATGAAAAGTTTATTTCTTCTCCACATCAGCTTTTCTTCTGCGTTTCA 300
DB 78 TTTTGTGAGTTTGGTAATTTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 25

RESULT 9

US-10-239-676-185
Sequence 185, Application US/10239676
Publication No. US20030082609A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
FILE REFERENCE: 5013.1003
CURRENT APPLICATION NUMBER: US/10/239,676
CURRENT FILING DATE: 2002-09-24
PRIOR APPLICATION NUMBER: PCT/EP01/03968

DE 10019058.8
DE 10019173.8
DE 10032529.7
DE 10043826.1
PRIOR FILING DATE: 2001-04-06
2000-04-06
2000-04-07
2000-06-30
2000-09-01
NUMBER OF SEQ ID NOS: 228
SEQ ID NO 185

LENGTH: 7657
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-185

Query Match 15.7%; Score 47.2; DB 14; Length 7657;
Best Local Similarity 47.6%; Pred. No. 0.33;
Matches 139; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 7 TTTTGTGAGTTTGGTAATTTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 66
DB 5513 TTTTGTGAGTTTGGTAATTTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 5572
QY 67 GAACCTTCAAGCTGGGAAGGCTGCATGAAATTCAGTTGCTCCAGACAGCTTC 126
DB 5573 TATTTTGTGAGTTTGGTAATTTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 5632
QY 127 GGGCGATCCTTTGGTAAGGGGCTCCGCGCTTTAAATTTCTTCTTCTACAGT 186
DB 5633 TTTTGTGAGTTTGGTAATTTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 5692
QY 187 CTTTGTGAGTTTGGTAATTTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 246
DB 5693 TTTTGTGAGTTTGGTAATTTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 5752
QY 247 ATCTGATGAAAAGTTTATTTCTTCTCCACATCAGCTTTTCTTCTGCGTTTCA 298
DB 5753 TTTTGTGAGTTTGGTAATTTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 5804

RESULT 10

US-10-311-455-1995
Sequence 1995, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining
TITLE OF INVENTION: Cytosine methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 1995
LENGTH: 7657
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1995

Query Match 15.7%; Score 47.2; DB 14; Length 7657;
Best Local Similarity 47.6%; Pred. No. 0.33;

```

APPLICANT: Vockley, Joseph G.
TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
FILE REFERENCE: 44921-5026
CURRENT APPLICATION NUMBER: US/10/240,425
CURRENT FILING DATE: 2002-09-30
PRIOR APPLICATION NUMBER: PCT/US01/09847
PRIOR FILING DATE: 2001-03-28
PRIOR APPLICATION NUMBER: US 60/193,446
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 1588
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1099
LENGTH: 201143
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20040033502A1 AL022318
US-10-240-425-1099

Query Match      15.5%; Score 46.8; DB 12; Length 201143;
Best Local Similarity 56.1%; Pred. No. 1.4;
Matches 87; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY      3  TTCCCTTTTTTTTTTTTGAAGATTATTTTACTGCTGGAATCTTTTGTAGAAACAC 62
Db      147983  TTCTCCTGTTTTTTTTTCTCAGATTGTGTACTGTGTGAAAACTTTGTACAGTGAT 148042

QY      63  GAAAGACCTTTCAAAGCCTGGAGAGGCTGCATGAAATTCAGTTGCTCTCCAGACAG 122
Db      148043  GAGAGACCATTCAGAGCCTTGGAGGAGCTACAAACCACTTTGACATTCGAAAAAGAG 148102

QY      123  CTTGCGGCATCTTTTGTAGGGGCTTCTCG 156
Db      148103  CTACGGAGATCTCCAGTGAGGGGTCTCCCTGG 148136

RESULT 13
US-10-424-599-36643/C
; Sequence 36643; Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 36643
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_133091C.1
US-10-424-599-36643

Query Match      15.5%; Score 46.6; DB 12; Length 351;
Best Local Similarity 56.1%; Pred. No. 0.14;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY      142  TAAAGGAGCTCCCGGCTTTTAAATTTCTTCTTCTCTACAGCTTTTGGAGTTTC 201
Db      184  TCGGGTTTTCGCAATTTTTTTTGTTTTTTTGTTCGCGATTATTTATTTTTT 125

QY      202  GATATTTCTTAAATTTCTTATGTTCATCACTCTGAGTTTTCATCTGATGAAAACT 261
Db      124  TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTT 65

QY      262  TATTTCTCTCCACATCAGCTTTTCTCTGCTGTT 298

```



```

/ STATE: MO
/ COUNTRY: USA
/ ZIP: 63105
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/897,126
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/520,678
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Henderson, Melodie W.
/ REGISTRATION NUMBER: 37,848
/ REFERENCE/DOCKET NUMBER: 6029-6836
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 314-727-5188
/ TELEFAX: 314-727-6092
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 30:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 270 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-08-897-126-30

Query Match
Best Local Similarity 13.7%; Score 41.2; DB 3; Length 270;
Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 142 TAAGGGCTTCCTGCTTTAAATTTCTTCTCTACAGTCTTTTGAGATTTC 201
DB 107 TAGGGCTCTCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 166
QY 202 GATAATTTCTTAATTTCTTAATGTTCAATCAGTCTGAGTTTCAATGATAACTT 261
DB 167 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 226
QY 262 TATTTT 267
DB 227 TTTTTC 232

RESULT 8
US-09-601-537-10/c
/ Sequence 10, Application US/09601537
/ Patent No. 6582691
/ GENERAL INFORMATION:
/ APPLICANT: Gallert, Karl-Christian
/ APPLICANT: Muller, Stefan
/ APPLICANT: Huls, Christoph
/ APPLICANT: Bohmisch, Britta
/ TITLE OF INVENTION: Expression Vector for the Production of
/ FILE REFERENCE: 50186/002001
/ CURRENT APPLICATION NUMBER: US/09/601,537
/ CURRENT FILING DATE: 2000-08-02
/ PRIOR APPLICATION NUMBER: PCT/EP99/00829
/ PRIOR FILING DATE: 1999-02-09
/ PRIOR APPLICATION NUMBER: DE 198 05 781.4
/ PRIOR FILING DATE: 1998-02-12
/ NUMBER OF SEQ ID NOS: 29
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 10
/ LENGTH: 441
/ TYPE: DNA
/ ORGANISM: Artificial Sequence

```

```

/ FEATURE:
/ OTHER INFORMATION: p153-NT3
/ US-09-601-537-10

Query Match
Best Local Similarity 13.6%; Score 40.8; DB 4; Length 441;
Matches 75; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 158 TTTTAAATTTCTTCTTCTCTACAGTCTTTTGGAGTTGTAATTTCTTAATT 217
DB 425 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 366
QY 218 TTTCTATGTTCAATCAGTCTGAGTTTCAATGAGAACTTATTTCTCTCCACT 277
DB 365 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 306
QY 278 CAGCTTTTCTT 289
DB 305 AATTAATTAATT 294

RESULT 9
US-09-800-729-78/c
/ Sequence 78, Application US/09800729
/ Patent No. 6605592
/ GENERAL INFORMATION:
/ APPLICANT: NI et al.
/ TITLE OF INVENTION: 32 Human secreted proteins
/ FILE REFERENCE: P2044P1
/ CURRENT APPLICATION NUMBER: US/09/800,729
/ CURRENT FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: PCT/US00/26013
/ PRIOR FILING DATE: 2000-09-22
/ PRIOR APPLICATION NUMBER: 60/155,709
/ PRIOR FILING DATE: 1999-09-24
/ NUMBER OF SEQ ID NOS: 217
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO: 78
/ LENGTH: 1141
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-800-729-78

Query Match
Best Local Similarity 13.6%; Score 40.8; DB 4; Length 1141;
Matches 63; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 158 TTTTAAATTTCTTCTTCTCTACAGTCTTTTGGAGTTGTAATTTCTTAATT 217
DB 1131 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1072
QY 218 TTTCTATGTTCAATCAGTCTGAGTTTCAATGAGAACTTATTTCTCTCCACT 273
DB 1071 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1016

RESULT 10
US-09-601-537-9/c
/ Sequence 9, Application US/09601537
/ Patent No. 6582691
/ GENERAL INFORMATION:
/ APPLICANT: Gallert, Karl-Christian
/ APPLICANT: Muller, Stefan
/ APPLICANT: Huls, Christoph
/ APPLICANT: Bohmisch, Britta
/ TITLE OF INVENTION: Expression Vector for the Production of
/ FILE REFERENCE: 50186/002001
/ CURRENT APPLICATION NUMBER: US/09/601,537
/ CURRENT FILING DATE: 2000-08-02
/ PRIOR APPLICATION NUMBER: PCT/EP99/00829
/ PRIOR FILING DATE: 1999-02-09
/ PRIOR APPLICATION NUMBER: DE 198 05 781.4

```

```

? PRIOR FILING DATE: 1998-02-12
? NUMBER OF SEQ ID NOS: 29
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 9
? LENGTH: 4121
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-601-537-9

```

| | | | | |
|--------------------------|-------|---------------|----------|-------------|
| Query Match | 13.6% | Score 40.8 | DB 4 | Length 4121 |
| Best Local Similarity | 56.8% | Pred. No. 0.2 | | |
| Matches 75; Conservative | 0 | Mismatches 57 | Indels 0 | Gaps 0 |

| | | | |
|----|------|--|------|
| OY | 158 | TTTTAAATTTCTTCTTCACAGCTTTTGAGTGGCAATTCTTAAT | 217 |
| Db | 4105 | TT | 4046 |
| OY | 218 | TTCATAATGCAATCAGCTGCATCATGAAACTTATTTCTCCACAT | 277 |
| Db | 4045 | TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGAATTTTGTGTTACAAT | 3988 |
| OY | 278 | CAGCTTTTCTT | 289 |
| Db | 3985 | AATAAATTAAT | 3974 |

| | | | | |
|-----------------------|-----------------|-----------------|-----------|---------------|
| Query Match | 13.4%; | Score 40.2; | DB 3; | Length 12980; |
| Best Local Similarity | 53.5%; | Pred. No. 0.39; | | |
| Matches 84; | Conservative 0; | Mismatches 73; | Indels 0; | Gaps 0 |

QY 94 ATGAAATATAGTTCGCTCTCCACGAAGAGCTCGGCGCATCTTTTGGTAAAGGGAGCTCC 153
 Db 9374 ATGAAGGTGGGGTAAACATCCGGGCTCTTAGCCATTCTCTGTTTTTTTTTTTTTTT 9433
 QY 154 TCGCTTTTAAATTTCTTTCTTCTCCACAGCTTTTGGAGTTTCGTATTTCTTA 213
 Db 9434 TTCCCTT 9493
 QY 214 TATTTCATATGTCATACATCTCAGTTTTCATCT 250
 Db 9494 TTTTTTTTTTTTTTTCTTCCCTCTTTTTCCCTT 9530

| | Query Match | Similarity | Score | DB 4 | Length | 12980 |
|----|-------------|---|--------------|------|------------|-------|
| | Best Local | 53.5% | Pred. No. | 0.3% | | |
| | Matches | 84 | Conservative | 0 | Mismatches | 73 |
| | | | | | Indels | 0 |
| | | | | | Gaps | 0 |
| Qy | 94 | ATGAATAATTCAGTTCGTCTCTCCAGAGAGCTTCGGCCATCTTTTGGTAAGGGGCTTCC | 153 | | | |
| Db | 9374 | ATGAAGGTTCGGGGTAACATCCCGCCTCTTAGCCATTTCTTTTTTTTTTTTTTTTT | 94338 | | | |
| Qy | 154 | TCGCTTTTAAATTTCTTTCTTCTCTACAGCTTTTTCGAGTTTCGTAATTTCTTA | 213 | | | |
| Db | 9434 | TT | 94938 | | | |
| Qy | 214 | TATTTTCTTAATGTTCAATACATCTAGTTTAACT | 250 | | | |
| Db | 9494 | TTTTTTTTTTTTTTTTTCTTTCTCTTTTCTTTTCTTT | 9530 | | | |

RESULT 13
US-09-621-976-1322/c
; Sequence 1322, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1322
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..249
US-09-621-976-1322

Query Match 13.3%; Score 40; DB 4; Length 249;
Best Local Similarity 57.0%; Pred. No. 0.14;
Matches 73; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 158 TTTTAAATTTCTTTCTTCTCTACAGCTTTTGGAGTTTGGTAATTTCTTAAT 217
DB 247 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 188
QY 218 TTTCTTTTGTTCATCAGCTCTGAGTTTTCATCTGATGAAATTTATTTCTCCACAT 277
DB 187 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 128
QY 278 CAGCTTTT 285
DB 127 CAATTTT 120

RESULT 14
US-09-014-969-14/c
; Sequence 14, Application US/09014969
; Patent No. 5965397
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavalley, Edward R.
; APPLICANT: Racie, Lisa A.
; APPLICANT: Werberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Agostino, Michael J.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/014,969
; FILING DATE:

CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-014-969-14

Query Match 13.3%; Score 40; DB 2; Length 2447;
Best Local Similarity 59.8%; Pred. No. 0.27;
Matches 67; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 158 TTTTAAATTTCTTTCTTCTCTACAGCTTTTGGAGTTTGGTAATTTCTTAAT 217
DB 2335 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2276
QY 218 TTTCTTTTGTTCATCAGCTCTGAGTTTTCATCTGATGAAATTTATTTCTC 269
DB 2275 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2224

RESULT 15
US-10-204-708-23
; Sequence 23, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: Olex, Alexander
; APPLICANT: Pfenberg, Kurt
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 23
; LENGTH: 11049
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-23

Query Match 13.3%; Score 40; DB 4; Length 11049;
Best Local Similarity 53.9%; Pred. No. 0.42;
Matches 82; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 101 TTGATTCGTCTCTCAGACAGCTTGGGCAATCTTTGGTAGGAGCTTCTCGCTT 160
DB 7020 TTTAATTTTATTTTAGAAGATTGGTTTTTGGTATGTAGATTAAGTAGATTT 7079
QY 161 TTTAATTTCTTTCTTTCTCTACAGCTTTTGGAGTTTGGTAATTTCTTAATTTTC 220
DB 7080 TTGAATTTTCTTTTATTTTATTTAGTGTTTTGTGTTTGTGTTGTTGTTGTTT 7139

Qy 221 TTATGTGATGACGCTGAGTTTCATCTGA 252
Db 7140 TTAAATTATTTAGTTGAGTTTGAGTTA 7171

Search completed: March 13, 2004, 00:51:04
Job time : 25.3958 secs

CC The present invention describes an activation-induced cytidine deaminase
CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
CC cytidine activity similar to APOBEC-1. AID has anti-allergic, anti-nausea,
CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
CC and can be used in gene therapy. AID polynucleotides are useful in
CC methods for identifying drugs for the treatment of B cell associated
CC immune system disorders, immunodeficiency diseases and allergies, such as
CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
CC drug allergy, allergic rhinitis, Rosen disease, DiGeorge disease, ataxia
CC telangiectasia, common variable immunodeficiency disorder, MHC (major
CC histocompatibility class) II deficiency disease, AIDS (auto
CC immunodeficiency syndrome), elevated IGE disorder, and IGE subclass
CC selection disorder. The DNA sequences encoding AID may be used for gene
CC therapy and the antibodies to the AID protein may be used for diagnosis
CC and treatment of these disorders. The present sequence represents a
CC genomic DNA sequence of human AID
CC
XX

Sequence 6564 BP; 1909 A; 1358 C; 1383 G; 1914 T; 0 U; 0 Other;

Query Match 100.0%; Score 301; DB 3; Length 6564;

Best Local Similarity 100.0%; Pred. No. 5,3e-56; Mismatches 0; Indels 0; Gaps 0;

Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TTTTCCTTTTGTGAGATTATTTTACGCTGGAATCTTTGTGAAAAC 60

3132 TTTTCCTTTTGTGAGATTATTTTACGCTGGAATCTTTGTGAAAAC 3191

61 ACGAAGAACTTCAAGCCTGGAGGGGCTGCATGAAATTCAGTTCCTCCAGAC 120

3192 ACGAAGAACTTCAAGCCTGGAGGGGCTGCATGAAATTCAGTTCCTCCAGAC 3251

121 AGCTTCGGCGCATCCTTTGTAAGGGGCTCTCCCTTTTAAATTTCTTTCTC 180

3252 AGCTTCGGCGCATCCTTTGTAAGGGGCTCTCCCTTTTAAATTTCTTTCTC 3311

181 TACAGCTTTTGTGAGTTGTAATTTCTTATTTTCTTATTTGTAATCAGCTCA 240

3312 TACAGCTTTTGTGAGTTGTAATTTCTTATTTTCTTATTTGTAATCAGCTCA 3371

241 GTTTCATCTGATGAAACTTATTTCTCTCCACATCAGCTTTTCTGCTGTTCA 300

3372 GTTTCATCTGATGAAACTTATTTCTCTCCACATCAGCTTTTCTGCTGTTCA 3431

301 C 301

3432 C 3432

RESULT 2
AAC55339
ID AAC55339 standard; DNA; 11204 BP.

XX AAC55339;

XX 05-FEB-2001 (first entry)

XX Human activation-induced cytidine deaminase genomic DNA SEQ ID NO:35.

XX Activation-induced cytidine deaminase; AID; cytidine deaminase;
XX immune related disease; allergy; allergic disease; antiasthmatic;
XX antiasthmatic; antiasthmatic; ophthalmological; anti-HIV; dermatological;
XX gene therapy; B cell associated immune system disorder; food allergy;
XX immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
XX IGA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;
XX drug allergy; allergic rhinitis; Rosen disease; DiGeorge disease; AIDS;
XX ataxia telangiectasia; common variable immunodeficiency disorder;
XX major histocompatibility class II deficiency disease;
XX auto immunodeficiency syndrome; IGE subclass selection disorder; ds.

XX Homo sapiens.

XX WO200058480-A1.

XX 05-OCT-2000.

XX 28-MAR-2000; 2000MO-JP001918.

XX 29-MAR-1999; 99JP-00087192.

XX 24-JUN-1999; 99JP-00178999.

XX 27-DEC-1999; 99JP-00371382.

XX (NIBS) JAPAN TOBACCO INC.

XX (HONJ) HONJO T.

XX Honjo T, Muramatsu M;

XX WPI; 2000-611715/58.

XX Claim 17; Page 163-170; 174pp; Japanese.

XX The present invention describes an activation-induced cytidine deaminase

XX (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has

XX cytidine activity similar to APOBEC-1. AID has anti-allergic, anti-nausea,

XX antiasthmatic, ophthalmological, anti-HIV and dermatological activities,

XX and can be used in gene therapy. AID polynucleotides are useful in

XX methods for identifying drugs for the treatment of B cell associated

XX immune system disorders, immunodeficiency diseases and allergies, such as

XX immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-

XX globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,

XX drug allergy, allergic rhinitis, Rosen disease, DiGeorge disease, ataxia

XX telangiectasia, common variable immunodeficiency disorder, MHC (major

XX histocompatibility class) II deficiency disease, AIDS (auto

XX immunodeficiency syndrome), elevated IGE disorder, and IGE subclass

XX selection disorder. The DNA sequences encoding AID may be used for gene

XX therapy and the antibodies to the AID protein may be used for diagnosis

XX and treatment of these disorders. The present sequence represents a

XX genomic DNA sequence of human AID

XX Sequence 11204 BP; 3305 A; 2273 C; 2373 G; 3253 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 301; DB 3; Length 11204;

XX Best Local Similarity 100.0%; Pred. No. 5,6e-56; Mismatches 0; Indels 0; Gaps 0;

XX Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 TTTTCCTTTTGTGAGATTATTTTACGCTGGAATCTTTGTGAAAAC 60

8347 TTTTCCTTTTGTGAGATTATTTTACGCTGGAATCTTTGTGAAAAC 8406

61 ACGAAGAACTTCAAGCCTGGAGGGGCTGCATGAAATTCAGTTCCTCCAGAC 120

8407 ACGAAGAACTTCAAGCCTGGAGGGGCTGCATGAAATTCAGTTCCTCCAGAC 8466

121 AGCTTCGGCGCATCCTTTGTAAGGGGCTCTCCCTTTTAAATTTCTTTCTC 180

8467 AGCTTCGGCGCATCCTTTGTAAGGGGCTCTCCCTTTTAAATTTCTTTCTC 8526

181 TACAGCTTTTGTGAGTTGTAATTTCTTATTTTCTTATTTGTAATCAGCTCA 240

8527 TACAGCTTTTGTGAGTTGTAATTTCTTATTTTCTTATTTGTAATCAGCTCA 8586

241 GTTTCATCTGATGAAACTTATTTCTCTCCACATCAGCTTTTCTGCTGTTCA 300

8587 GTTTCATCTGATGAAACTTATTTCTCTCCACATCAGCTTTTCTGCTGTTCA 8646

301 C 301

8647 C 8647

RESULT 3
ABS73286

ID ABS73286 standard; DNA; 11204 BP.
 AC ABS73286;
 XX
 DT 04-DEC-2002 (first entry)
 DE DNA encoding human translocation del(12p) protein #1.
 XX
 XX Chromosome aberration; oncogenic fusion protein; cancer;
 KM Proliferative disease; cellular protein isoform; heat shock protein 90;
 KM HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
 KM T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
 KM acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
 KM acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
 KM papillary thyroid carcinoma; Ewing's sarcoma; liposarcoma;
 KM rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200269900-A2.
 XX
 PD 12-SEP-2002.
 XX
 PF 01-MAR-2002; 2002WO-US006518.
 XX
 PR 01-MAR-2001; 2001US-0272751P.
 XX
 PA (CONF-) CONFORMA THERAPEUTICS CORP.
 XX
 PI Fritz LC, Burrows FJ;
 XX
 DR WPI; 2002-698710/75.
 DR P-PDB; ABG35082.
 XX
 PT Treating genetically-defined disease associated with chromosomal
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
 PT diseases, involves administering an inhibitor of heat shock protein 90.
 XX
 PS Disclosure; Page 242-245; 389pp; English.
 XX
 CC The invention describes a method of treating genetically-defined disease
 CC associated with chromosomal aberrations yielding oncogenic fusion
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous
 CC cell population, treating proliferative diseases associated with mutant
 CC protein or cellular protein isoforms (II) dependent on heat shock protein
 CC (HSP)-90, or selectively treating cells expressing (II) involving
 CC administering HSP90-inhibitor. The method is useful for treating
 CC genetically-defined disease with chromosomal aberration yielding
 CC oncogenic fusion protein, treating cancerous cells containing fusion
 CC protein in heterogeneous cell population, treating proliferative disease
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
 CC p53), or selectively treating cells expressing mutant protein or cellular
 CC protein isoform in a patient heterozygous for (II). The method is useful
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, AML, NHL and CMML,
 CC or a disease characterised by a solid tumour such as papillary thyroid
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
 CC synovial sarcoma. The method is also useful for treating viral
 CC infections. This represents the DNA sequence of a chromosome aberration
 XX
 SQ Sequence 11204 BP; 3305 A; 2273 C; 2373 G; 3253 T; 0 U; 0 Other;
 Query Match 100.0%; Score 301; DB 6; Length 11204;
 Best Local Similarity 100.0%; Pred. No. 5.6e-56;
 Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTTTCTTTTCTTTTCTTTTGAAGATTATTTTACGCTGGAATACCTTTGTGAAAACC 60
 DB 8347 TTTTCTTTTCTTTTCTTTTGAAGATTATTTTCTGGAATACCTTTGTGAAAACC 8406
 QY 61 ACGAAAGAACTTCAAGCGCTGGAAGCGCTGGAATACCTTTGTGAAAACC 120

DB 8407 ACGAAAGAACTTCAAGCGCTGGAAGCGCTGGAATACCTTTGTGAAAACC 8466
 QY 121 AGCTTGCGCGCAGCTTTTGGTAAGGCGCTTCTGCTTTTAAATTTCTTTCTTC 180
 DB 8467 AGCTTGCGCGCAGCTTTTGGTAAGGCGCTTCTGCTTTTAAATTTCTTTCTTC 8526
 QY 181 TACAGCTTTTGGAGATTCGTATATTTCTTATTTCTTATTTGTTCAATCACTCTCA 240
 DB 8527 TACAGCTTTTGGAGATTCGTATATTTCTTATTTCTTATTTGTTCAATCACTCTCA 8586
 QY 241 GTTTTCATCTGATGAAACTTTATTTCTCTCCACATCAGCTTTTCTTCTGCTGTTTCA 300
 DB 8587 GTTTTCATCTGATGAAACTTTATTTCTCTCCACATCAGCTTTTCTTCTGCTGTTTCA 8646
 QY 301 C 301
 DB 8647 C 8647
 RESULT 4
 ADB88952
 ID ADB88952 standard; DNA; 597 BP.
 XX
 AC ADB88952;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE AID gene as substrate used in AID mutagenic method.
 XX
 XX ds; cytosstatic; virucide; antidote; mutation;
 KW activation-induced cytidine deaminase; cell phenotype;
 KW monoclonal antibody; antigen; cross-reactivity; tumour;
 KW germ warfare agent; toxin; V region.
 XX
 OS Homo sapiens.
 XX
 XX Key
 FH mutation
 FT
 FT /tag= a
 FT /note= "mutation in P1-5 hybridoma cell"
 FT replace(12,A)
 FT /tag= b
 FT /note= "mutation in Ramos cell"
 FT replace(14,C)
 FT /tag= c
 FT /note= "mutation in CHO cell"
 FT replace(61,T)
 FT /tag= d
 FT /note= "mutation in P1-5 hybridoma cell"
 FT replace(66,A)
 FT /tag= e
 FT /note= "mutation in P1-5 hybridoma cell"
 FT replace(73,T)
 FT /tag= f
 FT /note= "mutation in CHO cell"
 FT replace(79,G)
 FT /tag= g
 FT /note= "mutation in Ramos cell"
 FT replace(85,G)
 FT /tag= h
 FT /note= "mutation in P1-5 hybridoma cell"
 FT replace(112,C)
 FT /tag= i
 FT /note= "mutation in Ramos cell"
 FT replace(118,G)
 FT /tag= j
 FT /note= "mutation in P1-5 hybridoma cell"
 FT replace(156,T)
 FT /tag= k
 FT /note= "mutation in P1-5 hybridoma cell"
 FT replace(161,A)
 FT /tag= l
 FT /note= "mutation in P1-5 hybridoma cell"

```

FT mutation replace(165,T,G)
FT /*tag= m
FT /note= "G mutation in Ramos cell, T mutation in CHO cell"
FT replace(178,T)
FT /*tag= n
FT /note= "mutation in Ramos cell"
FT replace(206,G)
FT /*tag= o
FT /note= "mutation in CHO cell"
FT replace(209,C)
FT /*tag= p
FT /note= "mutation in Ramos cell"
FT replace(218,T)
FT /*tag= q
FT /note= "mutation in Ramos cell"
FT replace(250,C)
FT /*tag= r
FT /note= "mutation in Ramos cell"
FT replace(261,G)
FT /*tag= s
FT /note= "mutation in Ramos cell"
FT replace(289,A)
FT /*tag= t
FT /note= "mutation in Ramos cell"
FT replace(303,T)
FT /*tag= u
FT /note= "mutation in Ramos cell"
FT replace(329,T)
FT /*tag= v
FT /note= "mutation in Ramos cell"
FT replace(350,G,T)
FT /*tag= w
FT /note= "T mutation in Ramos cell, G mutation in CHO cell"
FT replace(356,A)
FT /*tag= x
FT /note= "A mutation in Ramos cell, A mutation in CHO cell"
FT replace(357,G)
FT /*tag= y
FT /note= "mutation in Ramos cell"
FT replace(367,T, A)
FT /*tag= z
FT /note= "mutation in PI-5 hybridoma cell"
FT replace(375,A)
FT /*tag= aa
FT /note= "mutation in CHO cell"
FT replace(402,A)
FT /*tag= ab
FT /note= "mutation in PI-5 hybridoma cell"
FT replace(410,T)
FT /*tag= ac
FT /note= "mutation in PI-5 hybridoma cell"
FT replace(418,T)
FT /*tag= ad
FT /note= "mutation in PI-5 hybridoma cell"
FT replace(420,T)
FT /*tag= ae
FT /note= "mutation in PI-5 hybridoma cell"
FT replace(421,A)
FT /*tag= af
FT /note= "mutation in CHO cell"
FT replace(430,G)
FT /*tag= ag
FT /note= "mutation in Ramos cell"
FT replace(440,A)
FT /*tag= ah
FT /note= "mutation in CHO cell"
FT replace(449,T)
FT /*tag= ai
FT /note= "mutation in PI-5 hybridoma cell"
FT replace(495,A)
FT /*tag= aj
FT /note= "mutation in CHO cell"
FT replace(517,C)

```

```

FT /*tag= ak
FT /note= "mutation in Ramos cell"
FT replace(526,T, A)
FT /*tag= al
FT /note= "T mutation in PI-5 hybridoma cell, A mutation in
FT Ramos cell"
FT replace(533,A)
FT /*tag= am
FT /note= "mutation in PI-5 hybridoma cell"
FT replace(538,A)
FT /*tag= an
FT /note= "mutation in CHO cell"
FT replace(549,A)
FT /*tag= ao
FT /note= "mutation in CHO cell"
FT replace(568,T)
FT /*tag= ap
FT /note= "mutation in Ramos cell"
FT replace(570,T)
FT /*tag= aq
FT /note= "mutation in CHO cell"

MO2003061363-A2.
31-JUL-2003.
15-JAN-2003; 2003WO-US001149.
17-JAN-2002; 2002US-0350269P.
(YESH ) UNIV YESHIVA EINSTEIN COLLEGE.
Martin A, Scharff MD;
XX WPI; 2003-747983/70.
XX
XX Induction of mutations in gene expressed in eukaryotic cell useful e.g.
XX to produce antibodies with greater affinity or specificity for antigens
XX e.g. human pathogens by expressing an activation-induced cytidine
XX deaminase gene in cell.
XX
XX Example 4; Fig 6; 73pp; English.
XX
XX The invention relates to the induction of mutations in a gene expressed
XX in a eukaryotic cell, where the gene is operably linked to a promoter and
XX within 2 kb of promoter, by expressing a transgenic activation-induced
XX cytidine deaminase (Aid) gene in the cell. The method is useful for
XX producing mutated genes and proteins, determining the effect of a
XX mutation on a protein or cell phenotype, and producing genes and proteins
XX with altered properties. It is especially useful for producing antibodies
XX with altered affinities or specificities for an antigen, or
XX increased/decreased cross-reactivity for a second antigen, e.g. to
XX produce antibodies with greater affinity/specificity useful
XX therapeutically to treat tumours, viruses such as Ebola and Lassa Fever
XX or against germ warfare agents, toxins (e.g. ricin). The method may
XX similarly be used to alter affinity, specificity or cross-reactivity of a
XX monoclonal antibody. This sequence represents the Aid gene used in the
XX method of the invention to induce mutations in the sequence.
XX
XX Sequence 597 BP; 128 A; 164 C; 155 G; 150 T; 0 U; 0 Other;
XX
XX Query Match 39.3%; Score 118.2; DB 8; Length 597;
XX Best Local Similarity 97.6%; Pred. No. 1.8e-16;
XX Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
OY 18 TTGGAAGTATTTTACTGCTGGAATACCTTTGTAGAAAACCGAAGAACTTTGAA 77
OY |||
OY 421 TTCAAGAATATTTTACTGCTGGAATACCTTTGTAGAAAACCGAAGAACTTTGAA 480
OY 78 GCGTGGGAAGGCGCGCAAGAAATTCAGTGTCTCCAGACAGCTTCGCGCATCCTT 137
OY |||
OY 481 GCGTGGGAAGGCGCGCAAGAAATTCAGTGTCTCCAGACAGCTTCGCGCATCCTT 540

```


| QY | 138 TTG 140 |
|---------------------------|---|
| Db | 541 TTG 543 |
| RESULT 5 | |
| ID | ABX05468 standard; cDNA, 1543 BP. |
| XX | ABX05468 |
| AC | ABX05468; |
| DT | 17-JAN-2003 (first entry) |
| XX | |
| DE | Human novel polynucleotide #483. |
| XX | |
| XX | Human; gene; ss; genetic disorder; gene mapping; medical imaging; cancer; |
| KW | neurodegenerative disorder; lymphoid cell disorder; osteoporosis; |
| KW | Parkinson's disease; Alzheimer's disease; bone degenerative disorder; |
| KW | osteoarthritis; periodontal disease; liver fibrosis; viral infection; |
| KW | fungal infection; bacterial infection; autoimmune disease; diabetes; |
| KW | atopic dermatitis. |
| OS | Homo sapiens. |
| XX | |
| XX | WO200274961-A1. |
| PN | |
| PD | 26-SEP-2002. |
| XX | |
| PF | 14-MAR-2002; 2002MO-US005109. |
| XX | |
| PR | 15-MAR-2001; 2001US-00810173. |
| XX | |
| PA | (HYSEQ-) HYSEQ INC. |
| PI | Yang YF, Zhou P, Goodrich R, Aegund V, Zhang J, Zhao QA, Ren F; |
| PI | Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M, |
| PI | Wehrman T, Wang J, Wang D, Drmanac RT; |
| DR | WPI; 2003-040556/03. |
| DR | P-PSDB; ABU00390. |
| PT | New isolated polypeptides and polynucleotides, useful for preventing, |
| PT | treating or ameliorating medical conditions, such as cancer, |
| PT | neurodegenerative disorders, lymphoid cell disorders, bone degenerative |
| PT | disorders, and infections. |
| XX | |
| PS | Claim 1, SEQ ID NO 483; 235dp; English. |
| XX | |
| XX | The invention relates to human polynucleotides and the polypeptides they |
| CC | encode. The polynucleotides and polypeptides are useful in diagnostics, |
| CC | for forensic, gene mapping, medical imaging, identification of mutations, |
| CC | responsible for genetic disorders or other traits, assessing biodiversity |
| CC | and producing many other types of data and products dependent on DNA and |
| CC | amino acid sequences. They are also useful for preventing, treating or |
| CC | ameliorating medical conditions, such as cancer, neurodegenerative |
| CC | disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell |
| CC | disorders, osteoporosis, osteoarthritis, bone degenerative disorders, |
| CC | periodontal disease, liver fibrosis, infections (e.g. viral, fungal or |
| CC | bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis). |
| CC | Sequences ABX04966-ABX05511 represent human polynucleotides of the |
| CC | invention. Note: The sequence data for this patent is not represented in |
| CC | the printed specification but is based on sequence information supplied |
| CC | by the European Patent Office |
| XX | |
| SO | Sequence 1543 BP; 428 A; 334 C; 352 G; 429 T; 0 U; 0 Other; |
| Query Match | 39.3%; Score 118.2; DB 7; Length 1543; |
| Best Local Similarity | 97.6%; Pred. No. 1.9e-16; |
| Matches 120; Conservative | 0; Mismatches 3; Indels 0; Gaps 0. |
| QY | 18 TTGGAAGATTATTTTACTGCTGGGAATATCTTTGTGAGAAACCAAGAAAGAACTTGCAA 77 |
| 498 | TTCAAGAGATATTTTACTGCTGGGAATATCTTTGTGAGAAACCAAGAAAGAACTTGCAA 557 |

| | | | |
|----------|---|--|-----|
| YY | 76 | GCCTGGGAGGAGGCTGCATGAAGAAATTCAGTTCGTCTCCAGACAGCTTGGGGCATCCCTT | 13 |
| DB | 558 | GCTCTGGGAAGGAGCTGCATGAAGAAATTCAGTTCGTCTCTCCAGACAGCTTGGGGCATCCCTT | 617 |
| OY | 138 | TTG 140 | |
| DB | 618 | TTG 620 | |
| RESULT 6 | | | |
| AB873287 | | | |
| ID | AB873287 | standard; DNA; 2791 BP. | |
| XX | | | |
| AC | AB873287; | | |
| XX | | | |
| DT | 04-DEC-2002 | (first entry) | |
| XX | | | |
| DE | DNA encoding human translocation del(12p) protein #2. | | |
| XX | | | |
| KW | Chromosome aberration; oncogenic fusion protein; cancer; | | |
| KW | proliferative disease; cellular protein isoform; heat shock protein 90; | | |
| KW | HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder; | | |
| KW | T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML; | | |
| KW | acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML; | | |
| KW | acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour; | | |
| KW | papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma; | | |
| KW | rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds. | | |
| XX | | | |
| OS | Homo sapiens. | | |
| XX | | | |
| FN | MO200269900-A2. | | |
| XX | | | |
| PD | 12-SEP-2002. | | |
| XX | | | |
| PF | 01-MAR-2002; 2002W0-US006518. | | |
| PR | 01-MAR-2001; 2001US-0272751P. | | |
| XX | | | |
| PA | (CONF-) CONFORMA THERAPEUTICS CORP. | | |
| PI | Fritz LC, Burrows FU; | | |
| DR | WPI: 2002-698710/75. | | |
| XX | P-PSDB; AB935083. | | |
| PT | Treating genetically-defined disease associated with chromosomal | | |
| PT | aberrations yielding oncogenic fusion proteins, e.g. cell proliferative | | |
| PT | diseases, involves administering an inhibitor of heat shock protein 90. | | |
| XX | | | |
| PS | Disclosure; Page 246-247, 389pp; English. | | |
| XX | | | |
| XX | The invention describes a method of treating genetically-defined disease | | |
| CC | associated with chromosomal aberrations yielding oncogenic fusion | | |
| CC | proteins (I), treating cancerous cells containing (I) in a heterogeneous | | |
| CC | cell population, treating proliferative diseases associated with mutant | | |
| CC | protein or cellular protein isoforms (II) dependent on heat shock protein | | |
| CC | (HSP)-90, or selectively treating cells expressing (II) involving | | |
| CC | administering HSP90-inhibitor. The method is useful for treating | | |
| CC | genetically-defined disease with chromosomal aberration yielding | | |
| CC | oncogenic fusion protein, treating cancerous cells containing fusion | | |
| CC | protein in heterogeneous cell population, treating proliferative disease | | |
| CC | (e.g. rheumatoid arthritis or cancer) associated with mutant protein or | | |
| CC | cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g. | | |
| CC | p53), or selectively treating cells expressing mutant protein or cellular | | |
| CC | protein isoform in a patient heterozygous for (II). The method is useful | | |
| CC | for treating a disease e.g. haematopoietic disorder such as T or B cell | | |
| CC | lymphoma, chronic myeloid leukaemia (CML), APL, ALL, AML, NHL and CMML, | | |
| CC | or a disease characterised by a solid tumour such as papillary thyroid | | |
| CC | carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and | | |
| CC | synovial sarcoma. The method is also useful for treating viral | | |
| XX | infections. This represents the DNA sequence of a chromosome aberration | | |
| XX | | | |

Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;
 Query Match 39.3%; Score 118.2; DB 6; Length 2791;
 Best Local Similarity 97.6%; Pred. No. 2.1e-16;
 Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

18 TTGGAAGATTATTTTACTGCTGGAATATCTTTGTGAGAAACCAAGAAAGAACTTCAA 77
 497 TTCAAAGATTATTTTACTGCTGGAATATCTTTGTGAGAAACCAAGAAAGAACTTCAA 556

DB 78 GCCTGGAGAGGCTGCATGAAATTCAGTTCCTCTCCAGACAGCTTGGGCACTCTT 137
 557 GCCTGGAGAGGCTGCATGAAATTCAGTTCCTCTCCAGACAGCTTGGGCACTCTT 616

QY 138 TTG 140
 617 TTG 619

DB 617 TTG 619

RESULT 7
 ABS73288
 ID ABS73288 standard; DNA; 2791 BP.
 AC ABS73288;
 XX
 DT 04-DEC-2002 (first entry)
 XX
 DE DNA encoding human translocation del(12p) protein #3.
 XX
 XX Chromosome aberration; oncogenic fusion protein; cancer;
 XX proliferative disease; cellular protein isoform; heat shock protein 90;
 XX HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
 XX T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
 XX acute myeloid leukaemia; ALL; chronic myelomonocytic leukaemia; CMML;
 XX acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
 XX papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
 XX rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200269900-A2.
 XX
 PD 12-SEP-2002.
 XX
 PF 01-MAR-2002; 2002WO-US006518.
 XX
 PR 01-MAR-2001; 2001US-0272751P.
 XX
 PA (CONF-) CONFORMA THERAPEUTICS CORP.
 XX
 PI Fritz LC, Burrows FJ;
 XX
 DR P-PSDB; ABG95084.
 XX
 PT Treating genetically-defined disease associated with chromosomal
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
 PT diseases, involves administering an inhibitor of heat shock protein 90.
 XX
 PS Disclosure; Page 248-249; 389pp; English.

p53), or selectively treating cells expressing mutant protein or cellular
 CC protein isoform in a patient heterozygous for (II). The method is useful
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, ANL, NHL and CMML,
 CC or a disease characterised by a solid tumour such as papillary thyroid
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
 CC synovial sarcoma. The method is also useful for treating viral
 CC infections. This represents the DNA sequence of a chromosome aberration
 CC
 XX
 SQ Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;
 Query Match 39.3%; Score 118.2; DB 6; Length 2791;
 Best Local Similarity 97.6%; Pred. No. 2.1e-16;
 Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

18 TTGGAAGATTATTTTACTGCTGGAATATCTTTGTGAGAAACCAAGAAAGAACTTCAA 77
 497 TTCAAAGATTATTTTACTGCTGGAATATCTTTGTGAGAAACCAAGAAAGAACTTCAA 556

DB 78 GCCTGGAGAGGCTGCATGAAATTCAGTTCCTCTCCAGACAGCTTGGGCACTCTT 137
 557 GCCTGGAGAGGCTGCATGAAATTCAGTTCCTCTCCAGACAGCTTGGGCACTCTT 616

QY 138 TTG 140
 617 TTG 619

DB 617 TTG 619

RESULT 8
 AAC55312
 ID AAC55312 standard; cDNA; 2818 BP.
 AC AAC55312;
 XX
 DT 05-FEB-2001 (first entry)
 XX
 DE Human activation-induced cytidine deaminase encoding cDNA SEQ ID NO.7.
 XX
 XX Activation-induced cytidine deaminase; AID; cytidine deaminase;
 XX immune related disease; allergy; allergic disease; anti-allergic;
 XX antineoplastic; antineoplastic; ophthalmological; anti-HIV; dermatological;
 XX gene therapy; B cell associated immune system disorder; food allergy;
 XX immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
 XX IGA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;
 XX drug allergy; allergic rhinitis; Rosen disease; Digorge disease; AIDS;
 XX ataxia telangiectasia; common variable immunodeficiency disorder;
 XX major histocompatibility class II deficiency disease;
 XX auto immunodeficiency syndrome; IgG subclass selection disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT CDS 80..676
 FT /*tag= a
 FT /product= "activation-induced cytidine deaminase"
 XX
 PN WO2000058480-A1.
 XX
 PD 05-OCT-2000.
 XX
 PF 28-MAR-2000; 2000WO-JP001918.
 XX
 PR 29-MAR-1999; 99JP-00087192.
 PR 24-JUN-1999; 99JP-00178999.
 PR 27-DEC-1999; 99JP-00371182.
 XX
 PA (NIBS) JAPAN TOBACCO INC.
 PA (HONJ) HONJO T.
 XX
 PI Honjo T, Muramatsu M;
 XX
 DR WPI; 2000-611715/58.
 DR P-PSDB; AAB24198.

XX Nucleic acid encoding activation induced cytidine deaminase, useful as a
PT target for drug development for immune-related diseases including
PT allergies.
PS Claim 3; Page 135-139; 174pp; Japanese.
XX
XX The present sequence encodes human activation-induced cytidine deaminase
CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
CC cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic,
CC antisthmatic, ophthalmological, anti-HIV and dermatological activities,
CC and can be used in gene therapy. AID polynucleotides are useful in
CC methods for identifying drugs for the treatment of B cell associated
CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
CC globulinaemia, atopic dermatitis, allergic colitis, asthma, food allergy,
CC drug allergy, allergic rhinitis, Rosen disease, Digorge disease, ataxia
CC telangiectasia, common variable immunodeficiency disorder, MHC (major
CC histocompatibility class) class II deficiency disease, AIDS (auto
CC immunodeficiency syndrome), elevated Ige disorder, and Igg subclass
CC selection disorder. The DNA sequences encoding AID may be used for gene
CC therapy and the antibodies to the AID protein may be used for diagnosis
CC and treatment of these disorders
XX
SQ Sequence 2818 BP; 868 A; 548 C; 626 G; 776 T; 0 U; 0 Other;
Query Match 39.3%; Score 118.2; DB 3; Length 2818;
Best Local Similarity 97.6%; Pred. No. 2.1e-16;
Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 18 TTGAAGATTATTTTACTGCTGGAATCTTTGTGAGAAACCCAGAAAGACTTTCAA 77
DB 500 TTGAAGATTATTTTACTGCTGGAATCTTTGTGAGAAACCCAGAAAGACTTTCAA 559
QY 78 GCTGGAGAGGCTGCATGAAATTCAGTTCCTCTCCAGACAGCTGGCGCATCCTT 137
DB 560 GCTGGAGAGGCTGCATGAAATTCAGTTCCTCTCCAGACAGCTGGCGCATCCTT 619
QY 138 TTG 140
DB 620 TTG 622

RESULT 9
AAC55318
ID AAC55318 standard; DNA; 116 BP.
XX
AC AAC55318;
XX
DT 05-FEB-2001 (first entry)
XX
DE Human activation-induced cytidine deaminase exon 4 SEQ ID NO:14.
XX
KM Activation-induced cytidine deaminase; AID; cytidine deaminase;
KM immune related disease; allergy; allergic disease; antiallergic;
KM antianemic; antisthmatic; ophthalmological; anti-HIV; dermatological;
KM gene therapy; B cell associated immune system disorder; food allergy;
KM immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
KM IGA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis;
KM drug allergy; allergic rhinitis; Rosen disease; Digorge disease; AIDS;
KM ataxia telangiectasia; common variable immunodeficiency disorder;
KM major histocompatibility class II deficiency disease;
KM auto immunodeficiency syndrome; Igg subclass selection disorder; de.
XX
XX Homo sapiens.
XX
PN WO200058480-A1.
XX
PD 05-OCT-2000.
XX
PF 28-MAR-2000; 2000WO-JP001918.
XX
PR 29-MAR-1999; 99JP-00087192.

PR 24-JUN-1999; 99JP-00178999.
PR 27-DEC-1999; 99JP-00371382.
XX
XX (NINSB) JAPAN TOBACCO INC.
PA (HONU/) HONJO T.
XX
PI Honjo T, Muramatsu M;
XX
DR WPI; 2000-611715/58.
XX
PT Nucleic acid encoding activation induced cytidine deaminase, useful as a
PT target for drug development for immune-related diseases including
PT allergies.
XX
XX Claim 18; Page 151; 174pp; Japanese.
XX
PS The present invention describes an activation-induced cytidine deaminase
CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
CC cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic,
CC antisthmatic, ophthalmological, anti-HIV and dermatological activities,
CC and can be used in gene therapy. AID polynucleotides are useful in
CC methods for identifying drugs for the treatment of B cell associated
CC immune system disorders, immunodeficiency diseases and allergies, such as
CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
CC globulinaemia, atopic dermatitis, allergic colitis, asthma, food allergy,
CC drug allergy, allergic rhinitis, Rosen disease, Digorge disease, ataxia
CC telangiectasia, common variable immunodeficiency disorder, MHC (major
CC histocompatibility class) class II deficiency disease, AIDS (auto
CC immunodeficiency syndrome), elevated Ige disorder, and Igg subclass
CC selection disorder. The DNA sequences encoding AID may be used for gene
CC therapy and the antibodies to the AID protein may be used for diagnosis
CC and treatment of these disorders. The present sequence represents the
CC exon 4 genomic DNA sequence of human AID
XX
SQ Sequence 116 BP; 32 A; 25 C; 24 G; 35 T; 0 U; 0 Other;
Query Match 38.5%; Score 116; DB 3; Length 116;
Best Local Similarity 100.0%; Pred. No. 4.4e-16;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25 ATATATTTTACTGCTGGAATCTTTGTGAGAAACCCAGAAAGACTTTCAAAGCCTGG 84
DB 1 ATATATTTTACTGCTGGAATCTTTGTGAGAAACCCAGAAAGACTTTCAAAGCCTGG 60
QY 85 AAGGCTGCATGAAATTCAGTTCCTCTCCAGACAGCTGGCGCATCCTTTG 140
DB 61 AAGGCTGCATGAAATTCAGTTCCTCTCCAGACAGCTGGCGCATCCTTTG 116

RESULT 10
AAC55307
ID AAC55307 standard; cDNA; 2440 BP.
XX
AC AAC55307;
XX
DT 05-FEB-2001 (first entry)
XX
DE Mouse activation-induced cytidine deaminase encoding cDNA SEQ ID NO:1.
XX
KM Activation-induced cytidine deaminase; AID; cytidine deaminase;
KM immune related disease; allergy; allergic disease; antiallergic;
KM antianemic; antisthmatic; ophthalmological; anti-HIV; dermatological;
KM gene therapy; B cell associated immune system disorder; food allergy;
KM immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
KM IGA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis;
KM drug allergy; allergic rhinitis; Rosen disease; Digorge disease; AIDS;
KM ataxia telangiectasia; common variable immunodeficiency disorder;
KM major histocompatibility class II deficiency disease;
KM auto immunodeficiency syndrome; Igg subclass selection disorder; ss.
XX
XX Mus musculus.
XX
PH Key Location/Qualifiers

```
FT CDS 93..689
FT /*tag= a
FT /product= "activation-induced cytidine deaminase"
XX XX
XX PN MO200058480-A1.
XX
XX PD 05-OCT-2000.
XX
XX PF 28-MAR-2000; 2000WO-JP001918.
XX
XX PR 29-MAR-1999; 99JP-00087192.
XX 24-JUN-1999; 99JP-00178999.
XX 27-DEC-1999; 99JP-00371382.
XX
XX PA (NIBS ) JAPAN TOBACCO INC.
XX PA (HONJ/) HONJO T.
XX
XX PI Honjo T, Muramatsu M;
XX
XX PS WPI; 2000-611715/58.
XX DR F-PSDB; AAB24197.
XX
XX PT Nucleic acid encoding activation induced cytidine deaminase, useful as a
XX PT target for drug development for immune-related diseases including
XX PT allergies.
XX
XX PS Claim 3; Page 126-130; 174pp; Japanese.
XX
XX CC The present sequence encodes mouse activation-induced cytidine deaminase
XX CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
XX CC cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic,
XX CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
XX CC and can be used in gene therapy. AID polynucleotides are useful in
XX CC methods for identifying drugs for the treatment of B cell associated
XX CC immune system disorders, immunodeficiency diseases and allergies, such as
XX CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
XX CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
XX CC drug allergy, allergic rhinitis, Rosen disease, DiGeorge disease, ataxia
XX CC telangiectasia, common variable immunodeficiency disorder, MHC (major
XX CC histocompatibility class II deficiency disease, AIDS (auto
XX CC immunodeficiency syndrome), elevated IgG disorder, and IgG subclass
XX CC selection disorder. The DNA sequences encoding AID may be used for gene
XX CC therapy and the antibodies to the AID protein may be used for diagnosis
XX CC and treatment of these disorders
XX
XX SQ Sequence 2440 BP; 706 A; 546 C; 551 G; 636 T; 0 U; 1 Other;
XX
XX Query Match 33.4%; Score 100.6; DB 3; Length 2440;
XX Best Local Similarity 88.6%; Pred. No. 1.3e-12;
XX Matches 109; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
XX
XX QY 18 TTGGAAGATATTTTACTGCTGGAATCTTTGTGAAACACAGAAAGACTTTCAA 77
XX DB 513 TTCAAAGCTATTTTACTGCTGGAATCATTTGTGAAAAATCGTGAAGAACTTTCAA 572
XX
XX QY 78 GGCTGGAGAGGCGTCATGAAATTCAGTTCCTCTCCAGACAGCTTCGGCGATCCTT 137
XX DB 573 GGCTGGAGAGGCGTCATGAAATTCCTCGGCTAACCAACAATTCGGCGATCCTT 632
XX
XX QY 138 TTG 140
XX |||
XX DB 633 TTG 635
XX
XX RESULT 11
XX AA187261/c
XX ID AA187261 standard; cDNA; 326 BP.
XX
XX AC AA187261;
XX
XX XX 06-NOV-2001 (first entry)
XX
XX DE Human polynucleotide SEQ ID NO 7321.
```

```
XX XX
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation; ss.
XX
XX OS Homo sapiens.
XX
XX PN MO200164835-A2.
XX
XX PD 07-SEP-2001.
XX
XX PF 26-FEB-2001; 2001WO-US004927.
XX
XX PR 28-FEB-2000; 2000US-00515126.
XX PR 18-MAY-2000; 2000US-00577409.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Tang YT, Liu C, Drmanac RT;
XX
XX XX WPI; 2001-514838/56.
XX DR P-PSDB; AA007330.
XX
XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX PT and treating e.g. leukemia, inflammation and immune disorders.
XX
XX PS Claim 1; SEQ ID NO 7321; 1399pp + Sequence Listing; English.
XX
XX CC The invention relates to human polynucleotides (AA179941-AA193841) and
XX CC the encoded proteins (AA000010-AA01910) that exhibit activity elating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activity/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pat_sequences
XX
XX SQ Sequence 326 BP; 204 A; 23 C; 42 G; 57 T; 0 U; 0 Other;
XX
XX Query Match 16.7%; Score 50.4; DB 4; Length 326;
XX Best Local Similarity 49.6%; Pred. No. 0.082;
XX Matches 129; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
XX
XX QY 11 TTTTGTGGAAGATTTTACTGCTGGAATCTTTGTGAAACACAGAAAGACTTTCAA 70
XX DB 278 TTTTGTGGAAGATTTTACTGCTGGAATCTTTGTGAAACACAGAAAGACTTTCAA 219
XX
XX QY 71 TTCAAAGCTGGAAGGCGTCATGAAATTCAGTTCCTCTCCAGACAGCTTCGGCG 130
XX DB 218 TTTCGGGAATTTTGTGGGCTTTTAAAAATTTTATTTTTCCTCCCGGGGGG 159
XX
XX QY 131 CATCCTTTGGTAAGGCGCTCCTCGCTTTAAATTTCTTCTTCCTACAGTCTT 190
XX DB 158 TTTTGTGGAAGATTTTACTGCTGGAATCTTTGTGAAACACAGAAAGACTTTCAA 99
XX
XX QY 191 TTGAGATTGTAATTTCTTATATTTCTATTTGTCATCAGCTCAGTTCATCT 250
XX DB 98 TTTTGTGGAAGATTTTACTGCTGGAATCTTTGTGAAACACAGAAAGACTTTCAA 39
XX
XX QY 251 GATGAAACTTATTTCTCC 270
XX DB 38 AATTAATAGTTTATTTATTC 19
XX
XX RESULT 12
XX ABV55592/c
XX ID ABV55592 standard; cDNA; 490 BP.
```

XX ABV55592;
AC
XX
XX 17-SEP-2002 (first entry)
DT
XX
XX Human prostate expression marker cDNA 55583.
DE
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX MO200160860-A2.
PN
XX
XX 23-AUG-2001.
PD
XX
XX 20-FEB-2001, 2001MO-US005171.
PF
XX
XX 17-FEB-2000, 2000US-018319P.
PR 16-MAR-2000, 2000US-0189862P.
PR 25-MAY-2000, 2000US-0207454P.
PR 09-JUN-2000, 2000US-0211314P.
PR 18-JUL-2000, 2000US-0219007P.
PR 13-DEC-2000, 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan UB;
PI
XX
XX WPI, 2001-662795/76.
DR
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 10735, 11750P; English.
XX
XX The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement; (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 490 BP; 240 A; 122 C; 56 G; 70 T; 0 U; 2 Other;
SQ
Query Match 16.2%; Score 48.8; DB 5; Length 490;
Best Local Similarity 50.4%; Pred. No. 0.19;
Matches 119; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
QY 66 AGAAGCTTTCAAGGCTGGGAGGCTGATGAATAATTCAGTTCCTCCAGACAGCTT 125
DB 372 AGGGTTTTTTTAAACCGGGGAGAGGGGGGGGAGGAGGAGGAGGAGGAGGAGG 313
QY 126 GGGCGATCCTTTTGGTAGGGGCTCCCGCTTTTAAATTTCTTTCTTCTCTAAG 185
DB 312 GGGGTTCTCCGGGGGGGGGGGGGGTTTTTTTTTTTAAAGTTGGTTTTTTTTTTT 253
QY 186 TCTTTTGGAGTTTGGTATTTCTTATTTCTTATTTCTTATTTCTTATTTCTTATTT 245
DB 252 TT 193
QY 246 CATCTGATGAAAACCTTATTTCTCTCCATCAGCTTTTCTTCTGCTGTTTCA 301
DB 192 TTCA 137

RESULT 13
ID ADB54224 standard; DNA; 4316 BP.
XX
XX ADB54224;
AC
XX
XX 04-DEC-2003 (first entry)
DT
XX
XX Pretreated genomic DNA region 148.
DE
XX
XX colon cell proliferative disorder; non methylated CpG dinucleotide;
KW cytostatic; cancer; adenoma; carcinoma; cytosine methylation state; ds.
XX
XX Unidentified.
OS
XX
XX MO2003072821-A2.
PN
XX
XX 04-SEP-2003.
PD
XX
XX 27-FEB-2003, 2003MO-EP002035.
PF
XX
XX 27-FEB-2002, 2002EP-00004551.
PR
XX
XX (EPIC-) EPIGENOMICS AG.
PA
XX
XX Adorjan P, Burger M, Maier S, Nimrich I, Becker E, Lesche R;
PI Rujan T, Schmitt A;
XX
XX WPI, 2003-731620/69.
DR
XX
XX Detecting and differentiating between colon cell proliferative disorders
PT associated with a gene or its regulatory regions comprises contacting a
PT target nucleic acid in a biological sample obtained from the subject with
PT a reagent.
XX
XX Claim 32; SEQ ID NO 280; 74P; English.
XX
XX The invention relates to a novel method for detecting and differentiating
CC between colon cell proliferative disorders associated with at least one
CC gene or its regulatory regions. The method comprises contacting a target
CC nucleic acid in a biological sample obtained from the subject with at
CC least one reagent or a series of reagents, where the reagent or series of
CC reagents distinguishes between methylated and non methylated CpG
CC dinucleotides within the target nucleic acid. The molecules of the
CC invention demonstrate cytostatic activity whilst the method may useful
CC for detecting and differentiating between colon cell proliferative
CC disorders, including cancers such as colon adenoma and colon carcinoma.
CC The PNA (peptide nucleic acid)-oligonucleotides are useful as probes for
CC determining cytosine methylation state or single nucleotide
CC polymorphisms. The current sequence is that of the pretreated genomic DNA
CC region of the invention. This sequence is not shown within the
CC specification but is taken from Wipoweb.
XX
XX Sequence 4316 BP; 934 A; 0 C; 1090 G; 2292 T; 0 U; 0 Other;
SQ
Query Match 16.1%; Score 48.4; DB 9; Length 4316;
Best Local Similarity 47.7%; Pred. No. 0.3;
Matches 142; Conservative 0; Mismatches 156; Indels 0; Gaps 0;
QY 1 TTTTCTTTTCTTTTCTTTTGAAGATTTTACGCTGGAAATACCTTTGTGAAGAAC 60
DB 3890 TTTTAGTTTTTAAATTAATTTTGTGTTTTTATTTTATTTGATTTTGGATTTT 3949
QY 61 ACAGAAAGAACTTTCAAGGCTGGGAGGCTGATGAATAATTCAGTTCCTCCAGAC 120
DB 3950 TTGAAGAAAATTTATTTTAAGAGTTTGTGTATATATGATTTGTTTGAAGATGA 4009
QY 121 AGCTTGGGAGCATCTTTTGGTAGGGGCTCCCGCTTTTAAATTTCTTTCTTCTTC 180
DB 4010 AGGTTTAGATATAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 4069
QY 181 TACAGTCTTTTGGAGTTTGGTATTTCTTATATTTCTTATTTCTTCAACACCTCA 240

SQ Sequence 506 BP; 195 A; 101 C; 95 G; 107 T; 0 U; 8 Other;

Query Match 15.9%; Score 47.8; DB 8; Length 506;

Best Local Similarity 58.6%; Pred. No. 0.32; Mismatches 58; Indels 0; Gaps 0;

Matches 82; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 127 GGGGATCCCTTTTGGTAAGGGGCTCTGCTTTTAAATTTCTTTCTCTACAGT 186

Db 324 GGGCCTTTTATTTTAAAGGGGGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 265

QY 187 CTTTTTGAGTTTCGATATTTCTTAATTTCTTATGTGCACTCAGTTTC 246

Db 264 TT 205

QY 247 ATCTGATGAAGAACTTATTT 266

Db 204 GTTCCAATTAAGAACTTATTT 185

Search completed: March 12, 2004, 18:08:04
Job time : 123.53 secs

COMMENT

JAPAN TOBACCO INC, TASUKU HONJO
OS Homo sapiens (human)
PN JP 2001245669-A/8
PD 11-SEP-2001
PF 28-MAR-2000 JP 2000092981

PI TASUKU HONJO, MASAMICHI MURAMATSU
PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12,
PC A61P17/00,
PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC

C12N1/21,
PC C12N5/10, C12N9/78, C12P21/02, C12P21/08, C12N1/21, C12R1/19, PC
(C12N5/10, C12R1/91), C12N15/00, C12N5/00, C12N5/00, C12R1/91) CC
FH Key Location/Qualifiers.

FEATURES

SOURCE

1. 6564
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 301; DB 6; Length 6564;
Best Local Similarity 100.0%; Pred. No. 4,1e-50;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTCCTTTTCTTTTGAAGATTTTAACTGCTGGAADACTTTGTGAAAC 60
DB 3132 TTTTCCTTTTCTTTTGAAGATTTTAACTGCTGGAADACTTTGTGAAAC 3191
QY 61 ACGAAGAAGCTTTCAAGCCTGGAGAGGCTGCATGAAATTCAGTCTCTCCAGAC 120
DB 3192 ACGAAGAAGCTTTCAAGCCTGGAGAGGCTGCATGAAATTCAGTCTCTCCAGAC 3251
QY 121 AGCTTCGCGCATCTTTTGTGTAAGGGCTTCCTCGTTTAAATTTCTTTCTTC 180
DB 3252 AGCTTCGCGCATCTTTTGTGTAAGGGCTTCCTCGTTTAAATTTCTTTCTTC 3311
QY 181 TACAGCTTTTGTGAGTTTGTATATTTCTTATATTTCTTATGTTCAATCACTCTCA 240
DB 3312 TACAGCTTTTGTGAGTTTGTATATTTCTTATATTTCTTATGTTCAATCACTCTCA 3371
QY 241 GTTTTCATCTGATGAAACTTATTTCTCCCTCCACATCAGCTTTTCTTCTGCTTTCA 300
DB 3372 GTTTTCATCTGATGAAACTTATTTCTCCCTCCACATCAGCTTTTCTTCTGCTTTCA 3431
QY 301 C 301
DB 3432 C 3432

RESULT 2
BD016860 11204 bp DNA linear PAT 27-AUG-2002
LOCUS Novel cytidine deaminase.
DEFINITION BD016860
VERSION BD016860.1 GI:22558036
KEYWORDS JP 2001245669-A/33.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 11204)
AUTHORS Honjo, T. and Muramatsu, M.
TITLE Novel cytidine deaminase
JOURNAL Patent: JP 2001245669-A 33 11-SEP-2001;
JOURNAL JAPAN TOBACCO INC, TASUKU HONJO
OS Homo sapiens (human)
PN JP 2001245669-A/33
PD 11-SEP-2001
PF 28-MAR-2000 JP 2000092981
PI TASUKU HONJO, MASAMICHI MURAMATSU
PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12,
PC A61P17/00,
PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC

C12N1/21,
PC C12N5/10, C12N9/78, C12P21/02, C12P21/08, C12N1/21, C12R1/19, PC
(C12N5/10, C12R1/91), C12N15/00, C12N5/00, C12N5/00, C12R1/91) CC
FH Key Location/Qualifiers.

FEATURES

SOURCE

1. 11204
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 301; DB 6; Length 11204;
Best Local Similarity 100.0%; Pred. No. 3.7e-50;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTCCTTTTCTTTTGAAGATTTTAACTGCTGGAADACTTTGTGAAAC 60
DB 8347 TTTTCCTTTTCTTTTGAAGATTTTAACTGCTGGAADACTTTGTGAAAC 8406
QY 61 ACGAAGAAGCTTTCAAGCCTGGAGAGGCTGCATGAAATTCAGTCTCTCCAGAC 120
DB 8407 ACGAAGAAGCTTTCAAGCCTGGAGAGGCTGCATGAAATTCAGTCTCTCCAGAC 8466
QY 121 AGCTTCGCGCATCTTTTGTGTAAGGGCTTCCTCGTTTAAATTTCTTTCTTC 180
DB 8467 AGCTTCGCGCATCTTTTGTGTAAGGGCTTCCTCGTTTAAATTTCTTTCTTC 8526
QY 181 TACAGCTTTTGTGAGTTTGTATATTTCTTATATTTCTTATGTTCAATCACTCTCA 240
DB 8527 TACAGCTTTTGTGAGTTTGTATATTTCTTATATTTCTTATGTTCAATCACTCTCA 8586
QY 241 GTTTTCATCTGATGAAACTTATTTCTCCCTCCACATCAGCTTTTCTTCTGCTTTCA 300
DB 8587 GTTTTCATCTGATGAAACTTATTTCTCCCTCCACATCAGCTTTTCTTCTGCTTTCA 8646
QY 301 C 301
DB 8647 C 8647

RESULT 3
AB040430 11204 bp DNA linear PRI 03-OCT-2000
LOCUS Homo sapiens AID gene for activation-induced cytidine deaminase,
DEFINITION complete cds.
ACCESSION AB040430
VERSION AB040430.1 GI:9988407
KEYWORDS AID; activation-induced cytidine deaminase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (sites)
AUTHORS Muto, T., Muramatsu, M., Taniwaki, M., Kinoshita, K. and Honjo, T.
TITLE Isolation, tissue distribution, and chromosomal localization of the
human activation-induced cytidine deaminase (AID) gene
JOURNAL Genomics 68 (1), 85-88 (2000)
MEDLINE 20408890
PUBMED 10950930

REFERENCE 2 (sites)
AUTHORS Revy, P., Muto, T., Levy, Y., Geisemann, F., Plebani, A., Sanal, O.,
Catalan, N., Forveille, M., Dufourcq-Lageouse, R., Gennery, A.,
Tezcan, I., Erey, F., Kayserili, H., Ugazio, A.G., Brousse, N.,
Muramatsu, M., Notarangelo, L.D., Kinoshita, K., Honjo, T., Fischer, A.
and Durandy, A.

Activation-induced cytidine deaminase (AID) deficiency causes the
autosomal recessive form of the Hyper-IgM syndrome (HIGM2)
JOURNAL Cell 102 (5), 565-575 (2000)
MEDLINE 20460541
PUBMED 11007475
REFERENCE 3 (bases 1 to 11204)
AUTHORS Muto, T., Muramatsu, M., Taniwaki, M., Kinoshita, K. and Honjo, T.
TITLE Direct Submission

of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAR-REPORT.

FEATURES

Source

Location/Qualifiers
1. 71132
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-438L7"
complement(1..1448)
/note="overlaps bases 1..1448 of clone AC092490"
/function="clone overlap"
187..408
/standard_name="57233"
439..560
/standard_name="92005"
774..881
/rpt_family="MIR"
905..1190
/rpt_family="AluSg"
1191..1213
/rpt_family="AT_rich"
1591..1807
/standard_name="6612"
1744..1819
/standard_name="8198"
1966..2264
/rpt_family="AluSg"
3286..3328
/rpt_family="(TTTC)n"
4097..4249
/rpt_family="AluSg"
4360..4654
/rpt_family="AluY"
complement(5140..5262)
/rpt_family="FLAM_C"
5275..5304
/rpt_family="AT_rich"
5305..5563
/rpt_family="AluO"
5564..5599
/rpt_family="AT_rich"
5644..5733
/rpt_family="MSTD"
5734..6012
/rpt_family="AluSg"
6020..6336
/rpt_family="AluSx"
6337..6629

/rpt_family="MSTD"
6330..6650
/rpt_family="(CAA)n"
complement(6711..6817)
/rpt_family="MIR"
complement(7033..7327)
/rpt_family="AluSx"
7607..7745
/rpt_family="FLAM_A"
7980..8047
/rpt_family="MIR"
complement(8050..8549)
/rpt_family="LTR47A"
8567..8698
/rpt_family="MIR"
complement(8815..9101)
/rpt_family="AluSg"
complement(11227..12175)
/rpt_family="LTR5"
12625..12900
/standard_name="87432"
complement(13547..13664)
/rpt_family="LIMB5"
complement(13665..13979)
/rpt_family="AluY"
complement(13980..14193)
/rpt_family="LIMB5"
14622..14924
/rpt_family="AluY"
complement(15813..17136)
/rpt_family="LIPAB"
complement(17272..17395)
/rpt_family="FLAM_C"
17426..17575
/rpt_family="MIR"

Query Match

Best Local Similarity 98.9%; Score 297.8; DB 9; Length 71132;
Matches 299; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTTTCCTTTTCTTTTCTTTTGAAGATTATTTTACTGCTGGAATCTTTGTGAGAAAC 60
DB 42890 TTTTCCTTTTCTTTTCTTTTGAAGATTATTTTACTGCTGGAATCTTTGTGAGAAAC 42949
QY 61 ACGAAAGACTTTTAAAGCTCGGAGAGGCTGATGAAATTCAGTCCGCTCCAGAC 120
DB 42950 ACGAAAGACTTTTAAAGCTCGGAGAGGCTGATGAAATTCAGTCCGCTCCAGAC 43009
QY 121 AGCTTGCGGCGATCCTTTTGTAAGGGGCTTCGCTTTTAATTTCTTCTTCTC 180
DB 43010 AGCTTGCGGCGATCCTTTTGTAAGGGGCTTCGCTTTTAATTTCTTCTTCTC 43069
QY 181 TACAGCTTTTGTGAGTTTCGTATATTTCTTATTTTCTTATTTGTCATGACCTCA 240
DB 43070 TACAGCTTTTGTGAGTTTCGTATATTTCTTATTTTCTTATTTGTCATGACCTCA 43129
QY 241 GTTTTCATCTGATGAAGAACTTATTTCTCTCCACATCAGCTTTTCTTCTGCTTCA 300
DB 43130 GTTTTCATCTGATGAAGAACTTATTTCTCTCCACATCAGCTTTTCTTCTGCTTCA 43189
QY 301 C 301
DB 43190 C 43190

RESULT 5

HSAS77811 HSAS77811 335 bp mRNA linear PRI 22-JUL-2003
LOCUS Homo sapiens partial mRNA for activation-induced cytidine deaminase
DEFINITION (AID gene).
ACCESSION AJ577811
VERSION AJ577811.1 GI:33145978
KEYWORDS activation-induced cytidine deaminase; AID gene.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1
TAKHAR, P., COKE, H.A., FEAR, D., SMURTHWAITE, L., DURHAM, S.R. and
GOLD, H.J.
TITLE Allergen and activation-induced deaminase drive heavy-chain class switching to IgE in the nasal mucosa of hay fever patients
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 335)
AUTHORS Takhar, P
TITLE Direct Submission
JOURNAL Submitted (21-JUL-2003) Takhar P., Biomedical Sciences Randall Centre, King's College London, New Hunt House, Guy's Campus, St Thomas St, London, SE1 1UL, UNITED KINGDOM
FEATURES
source
1. 335
/organism="Homo sapiens"
/mol_type="mRNA"
/isolation_source="nasal biopsy from inferior turbinate of an allergic rhinitis patient"
/db_xref="taxon:9606"
/clone="NR02"
/sex="female"
/cell_type="B lymphocyte"
/tissue_type="nasal mucosa"
/dev_stage="adult"
/country="United Kingdom"
1. 335
/gene="AID"
/CDS
1. 335
/function="editing (involved in somatic hypermutation and class switch recombination)"
/evidence="experimental"
/product="activation-induced cytidine deaminase"
/protein_id="CAI2261.1"
/db_xref="GI:33145979"
/translation="DPERCYRVTFWTSNPPCDCAHVADFLRGNNPISLRIFTARLYFCBDRKAEPEGLRRLHRAVQIAIMTFKDYFCWTFVENHRTFAMGGLHNSVRLSRQRRILL"
ORIGIN
Query Match 39.8%; Score 119.8; DB 9; Length 335;
Best Local Similarity 98.4%; Pred. No. 5.6e-14;
Matches 121; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 18 TTGAAGATATTTTACGCTGGAATACCTTTGTAGAAAACGAGAAAGAACTTTCAA 77
DB 213 TTCAAGATATTTTACGCTGGAATACCTTTGTAGAAAACGAGAAAGAACTTTCAA 272
QY 78 GCCTGGAGAGGCTGCATGAATTCAGTTCCTCTCCAGACAGCTTCGGCGCATCTT 137
DB 273 GCCTGGAGAGGCTGCATGAATTCAGTTCCTCTCTCCAGACAGCTTCGGCGCATCTT 332
QY 138 TTG 140
DB 333 TTG 335
RESULT 6
AF529837 596 bp mRNA linear SYN 17-SEP-2002
LOCUS Mus musculus clone 10 transgenic Homo sapiens AID (AID) mRNA,
DEFINITION complete cds.
ACCESSION AF529837
VERSION AF529837.1 GI:22297261
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 596)
MARTIN, A. and SCHARFF, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
REFERENCE 2 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA
FEATURES
source
1. 596
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="10"
/cell_type="hybridoma PL-5"
/transgenic
1. 596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. 596
/gene="AID"
1. 596
/gene="AID"
/codon_start=1
/product="AID"
/protein_id="P44954.24.1"
/db_xref="GI:22297262"
/translation="MDSLLNMRKFLYQFNWVWAKGRRETYLCVYVRRDSATSPSLDFGLRNKGCHVELFLRYISDMDLPGCYRVTFWTSNPPCDCAHVADFLRGNNPISLRIFTARLYFCBDRKAEPEGLRRLHRAVQIAIMTFKDYFCWTFVENHRTFAMGGLHNSVRLSRQRRILLPLRYVDLDAFRTGL"
ORIGIN
Query Match 39.6%; Score 119.2; DB 12; Length 596;
Best Local Similarity 97.6%; Pred. No. 6.7e-14;
Matches 121; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 17 TTGAAGATATTTTACGCTGGAATACCTTTGTAGAAAACGAGAAAGAACTTTCAA 76
DB 420 TTCAAGATATTTTACGCTGGAATACCTTTGTAGAAAACGAGAAAGAACTTTCAA 479
QY 77 AGCTGGAGAGGCTGCATGAATTCAGTTCCTCTCCAGACAGCTTCGGCGCATCTT 136
DB 480 AGCTGGAGAGGCTGCATGAATTCAGTTCCTCTCTCCAGACAGCTTCGGCGCATCTT 539
QY 137 TTG 140
DB 540 TTG 543
RESULT 7
AC094826/c
AC094826 241757 bp DNA linear HTG 09-MAY-2003
LOCUS Rattus norvegicus clone CH230-2013, WORKING DRAFT SEQUENCE.
ACCESSION AC094826
VERSION AC094826.4 GI:30466489
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 241757)
MURPHY, D., MARIE, M., METZGER, M., LEE, A., ABRAHAMSON, S., ADAMS, C., ALDER, J., ALLEN, C., ALLEN, H., ALSTROCK, S., AMIN, A., ANGUSANO, D., ANYALEBECHI, V., AYOGLI, A., AYODEJI, M., BACA, E., BADEN, H., BALDWIN, D., BANDARANAIKE, D., BARBER, M., BARNSTEAD, M., BENAHMED, F., BIEWALD, K., BLAIR, J., BLANKENBURY, K., BLYTH, P., BROWN, M., BRYANT, N., BUHAY, C., BURCH, P., BURRELL, K., CALDERON, E., CARDENAS, V., CARTER, K., CAVAZOS, I., CEASAR, H., CENTER, A.,

Chacko, J., Chavre, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, L., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carrillo, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Dublin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, M., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, R., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kover, C., Kowals, C., Kratt, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, Y., Liu, J., Liu, Y., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenschew, M., Louised, H., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Mallory, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackeme, O., Okunolu, G., Olariu, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Plimack, C., Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L., L., Piazzi, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodery, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Saverly, G., Scherer, S., Scott, G., Shaheen, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C.D., Smales, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steele, M., Strong, R., Sutton, A., Swatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, D., Waldron, L., Walker, B., Wang, J., Wang, R., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wleczek, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhuesen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G., and Gibbs, R.A.

Unpublished
Direct Submission
2 (bases 1 to 241757)
Morley, K.C.

Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 241757)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:23101539.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/atlas/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>

CONTACT: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GBMC
Center clone name: CH230-2013
Summary Statistics
Assembly program: Atlas
Consensus quality: 227874 bases at least Q40
Consensus quality: 231202 bases at least Q30
Consensus quality: 233430 bases at least Q20
Estimated insert size: 245502; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 241757: contig of 241757 bp in length.
Location/Qualifiers
1. 241757
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-2013"
1. 3654
/note="wgs contig"
complement(233839). (239261)
/note="clone boundary
clone end:17
site:ECORI
end_sequence:BH289756"

ORIGIN
Query Match 39.4%; Score 118.6; DB 2; Length 241757;
Best Local Similarity 78.5%; Pred. No. 3e-14;
Matches 142; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

| QY | 1 | TTTTCTTTT | TTTTTTTGAAGATATTTTACTGTCGTAATACCTTTGTGAAGAC | 60 |
|----|--------|---|---|----|
| DB | 110257 | TTATGCAATTAATTTCTCTTAAAGCTATTTTACTGCTGGAATACATTTGTAAATATC | 110199 | |
| QY | 61 | ACGAAGAAGACTTCAAGAGCTGGGAAGGCTGCATGAATAATTCAGTCTCCAGAC | 120 | |
| DB | 110197 | ATGAAAGAACTTTCAGAGCTGGGAAGGCTGCATGAATAATTCAGTCTCCAGAC | 110138 | |
| QY | 121 | AGCTTCGGCAGCATCTTTTGTGAAGAGGCTTCCTGCTTTTAAATTTCTTTCTTC | 180 | |
| DB | 110137 | AGCTTCGGCAGCATCTTTTGTGAAGAGGCTTCCTGCTTCCTCTCTCTCTCTCTC | 110078 | |
| QY | 181 | T 181 | | |
| DB | 110077 | T 110077 | | |

RESULT 8
AC109119 239506 bp DNA linear HTG 08-OCT-2002
DEFINITION Rattus norvegicus clone CH230-98A4, *** SEQUENCING IN PROGRESS ***.
AC109119
AC109119.5 GI:22856767
VERSION
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 239506)

AUTHORS

Muzny, D., Matile, Metzcker, M., Lee, S., Abbramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Aleshcok, S., Altin, A., Arguiano, D., Aryalchekch, V., Aoyagi, A., Ayodeji, M., Bacé, E., Baden, H., Baldwin, D., Bandatanake, D., Barber, M., Barnesfeed, M., Benahmed, F., Biswal, K., Bhatt, J., Blankenbush, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Butch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cessari, H., Center, A., Chacko, J., Chavez, P., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., De Souza, L., Davila, M. L., Davis, C., Day, Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Desam, C., Ding, Y., Dinh, H., Diya, K., Draeger, H., Dugan-Tochas, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabris, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregesorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Hareley, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladin, S. L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenzshewal, L., Loulsgesed, H., Lorado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawung, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Min, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervys, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackeleleh, O., Okwomou, G., Olampunsaagon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poldrecker, A., Popovic, D., Primus, E., Pu, L., L., Puzio, M., Quiroz, J., Rachin, E., Reeves, K., Reiter, M. A., Reish, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojars, A., Rose, M., Rose, R., Riz, S. U., Sanders, W., Savary, G., Scheier, S., Scott, G., Shatman, S., Shen, H., Shetty, N., Shvartsbeyn, A., Sleson, I., Slitter, C. D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svetek, A., Taber, P., Taylor, C., Taylor, T., Thomas, K., Thomas, S., Tingey, A., Trejos, Z., Uman, K., Valas, R., Vera, V., Villanasa, D., Waliron, L., Walker, B., Wang, Q., Wang, S., Warren, U., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Neiderhause, G., Weiss, R. A., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G., and Gibbs, R. A.

TITLE
JOURNAL
REFERENCES
AUTHORS
TITLE
JOURNAL

Direct Submission
unpublished
2 (bases 1 to 259506)

REFERENCES
AUTHORS
TITLE
JOURNAL

Submitted (08-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Sep 14, 2002 this sequence version replaced g1:125338869. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.ngsc.bcm.tmc.edu/projects/rzt/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine

Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----

Center clone name: CH230-98A4

```
Assembly program: Pirat2 version 0.990339
Consensus quality: 225337 bases at least Q40
Consensus quality: 229008 bases at least Q30
Consensus quality: 230948 bases at least Q20
Estimated insert size: 248565; sum-of-contigs estimation
Quality coverage: 4x in Q20 base; sum-of-contigs estimation
```

```

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hbrc.bcm.tmc.edu/docs/genbank\_draft\_data.html)
* NOTE: This is a "working draft" sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 259506: contig of 259506 bp in length.

```

FEATURES

10

```
misc_feature    /organism="Rattus"
misc_feature    /mol_type="genomic"
misc_feature    /db_xref="taxon:10"
misc_feature    /clone="CH30-98A4"
misc_feature    1..1136
misc_feature    /note="wgs_contig"
misc_feature    253785..25514
misc_feature    /note="wgs_contig"
misc_feature    258349..259506
misc_feature    /note="wgs_contig"
```

ORIGIN

| | | | | |
|-----------------------|----------------|-----------------|----------|---------------|
| Query Match | 39.4% | Score 118.6 | DB 2 | Length 259506 |
| Best Local Similarity | 78.5% | Pred. No. 3e-14 | | |
| Matches 142 | Conservative 0 | Mismatches 39 | Indels 0 | Gaps 0 |

Db 20666 TTATGCATTATTTCTCTCTAAGACCTAATTTTACCTCGCIGGAATACATTTGAGAAATTC 20667

Dd 20606 ATGAAAGACTTTCAAGCTTGGAAAGGCTGCATGAATACCTCGTCAAGCTATACCATGAC 2031

D5 20546 AGCTTCGGCAICCTTTGGAAAGTCCACACACACACACACACACACAC

DB 20486 J 20486

AC120617/c
10CTIS
AC120617
345098 bp
DNA 1linear
HTG 15-NOV-2002

```

VERSION AC120617.4  GI:25008441

```

VERSION AC12061/.4 G1:25006441

KEYWORDS
SOURCE
ORGANISM

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FILLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS

1 (bases 1 to 345098)
Muzny,D,Marie, Metzker,M,lee, Abramson,S, Adams,C, Alder,J, Allen,C, Allen,R, Alsbrooks,S, Amin,A, Anguiano,D, Anyalebech,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Bismuth,K, Blair,J, Blankenship,K, Blyth,P, Brown,M, Bryant,N, Bunay,C, Burch,P, Burrell,K, Calderon,E, Cardenas,V, Carter,K, Cavazos,I, Caesar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denson,S, Deramo,C, Ding,Y, Dim,H, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Evans,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Frazer,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Gebregiorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W, Gunaratne,P, Haaland,M, Hamill,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,M, Hollins,B, Howells,S, Huylk,S, Hume,J, Idlerbird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowis,C, Kraft,C,L, Lebow,H, Leyan,J, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorenshewa,L, Louisedge,H, Lozano,R,J, Lu,X, Ma,J, Maheshwari,M, Mahindratne,M, Mahmoud,M, Malloy,K, Mangum,A, Mangun,B, Mapua,P, Martin,K, Martin,R, Martinez,E, Mawhinney,S, McLeod,M,P, McNeill,T,Z, Meenen,E, Mloaavljovic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Mundassa,N, Murphy,M, Natl,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nwaokwelen,O, Okunolu,G, Olamunsgoon,A, Pal,S, Parks,K, Pasternak,S, Paul,H, Perez,A, Perez,L, Plankoch,C, Plopper,F, Poindexter,A, Popovic,D, Primus,E, Pu,L, L, Pazo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R, Rellly,B, Rellly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rivers,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,M, Saverly,G, Scherer,S, Scott,G, Shatsman,S, Shen,H, Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Sma3,D, Sneed,A, Sodergren,E, Song,X-Z, Sorelle,R, Sosa,J, Steidle,M, Strong,R, Sutton,A, Svatek,A, Tabor,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Uemari,K, Valas,R, Vera,V, Villaseña,D, Waldron,L, Walker,B, Wang,J, Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Willson,R, Wleczky,R, Wooden,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von Niederhausen,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstock,G, and Gibbs,R,A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 345098)
Worley,K,C.

Direct Submission
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 345098)
Rat Genome Sequencing Consortium.

REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23265477.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GWGN
Center clone name: CH230-3215
----- Summary Statistics
Assembly program: PHRAP version 0.990329
Consensus quality: 19191 bases at least Q40
Consensus quality: 194508 bases at least Q30
Consensus quality: 195878 bases at least Q20
Estimated insert size: 227256; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 330403: contig of 330403 bp in length
* 330404 330503: gap of unknown length
* 330504 331603: contig of 1100 bp in length
* 331604 331703: gap of unknown length
* 331704 337738: contig of 6035 bp in length
* 337739 337838: gap of unknown length
* 337839 345098: contig of 7260 bp in length.

FEATURES
source

1..345098
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-3215"
/complement(1177532..178431)
/note="clone boundary
clone_end:3p6
site:
end_sequence:RH328475"
179290..181054
/note="wgs end extension
clone_end:3p6"

misc_feature

ORIGIN

Query Match 39.4%; Score 118.6; DB 2; Length 345098;
Best Local Similarity 78.5%; Pred. No. 2.8e-14;
Matches 142; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1 TTTTCCTTTTCTTTTCTTTTGAAGATTTTTCCTGCTGGAATCTTTGAGAAACC 60
DB 141574 TTATGCATTATTTCTCTCTTAAGACATTTTTCCTGGAATCAATTTGAGAAATC 141515
QY 61 ACGAAGAAGCTTTAAAGCCTGGAAGGCTGATGAATTCAGTTCGCTCCAGAC 120
DB 141514 ATGAAGAAGCTTTAAAGCCTGGAAGGCTGATGAATTCAGTTCGCTCCAGAC 141455
QY 121 AGCTTCGCGCATCTTTTGTGAAGGGGCTTCCTGCTTTTAAATTTCTTTTCTC 180

Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Cricetinae;
REFERENCE 1 (bases 1 to 591)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
MEDLINE 22222991
PUBMED 12202747
REFERENCE 2 (bases 1 to 591)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA
FEATURES
Source location/qualifiers
1..591
/organism="Cricetus griseus"
/mol_type="mRNA"
/db_xref="taxon:10029"
/clone="15"
/cell_line="CHO"
/transgenic
1..591
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
<1..591
/gene="AID"
<1..591
/gene="AID"
/codon_start=1
/product="AID"
/protein_id="AA055442.1"
/db_xref="gi:22297296"
/translation="SLMNRKFLYQFNKVMWAKRCETIYCIYVKKRDSATSFSLDF
GLYNNKQCHVEFLFLRYISDWDDPGRCYVMTFTMSPCYDARVADFRLGNPVL
SLRFLYALVYCEIDRKAEPEGLRLHAGVQIAIMTPDYCYCWNTEVENHERFKAM
EGHENSRYLSRQLRRLILPLYVEDDLRDARFTGR"
ORIGIN
Query Match 39.3%; Score 118.2; DB 12; Length 591;
Best Local Similarity 97.6%; Pred. No. 1.1e-13;
Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 18 TTGAAGATATTTTATCTGCTGGAATCTTTGTAGAAACCAAGAAAGACTTTCAA 77
Db 415 TTCAAAAGTTATTTTACTGCTGGAATCTTTGTAGAAACCAAGAAAGACTTTCAA 474
QY 78 GCCGCGGAAGGCGCATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCCTT 137
Db 475 GCCGCGGAAGGCGCATGAAATTCAGTTGCTCTCTCCAGACAGCTTCGGCGCATCCTT 534
QY 138 TTG 140
Db 535 TTG 537
RESULT 13
AF529816 596 bp mRNA linear PRI 17-SEP-2002
LOCUS AF529816
DEFINITION Homo sapiens clone Ramos 2 AID (AID) mRNA, partial cds.
ACCESSION AF529816
VERSION AF529816.1 GI:22297219
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
REFERENCE 2 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.

```

TITLE
JOURNAL
Direct Submission
Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA
FEATURES
source
1. 596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 2"
1. .596
/gene="AID"
1. .>596
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/gene="AID"
/codon_start=1
/product="AID"
/protein_id="AAM95403.1"
/db_xref="GI:22297220"
/translation="MDSLIMNRKFLYQFKNVWAKGRRELYCYVKKRDSATGSL
DFYLRNKGCHVELLFLRYISPMDDPQRCYVWTFSTRSPCYDCARVADFLRNP
NLSIRLETARLYFCEDRKAPPEGLRRLHSAVGQVIAIMTFKDYFCWNTFVENHERTEK
AMEGLHENSVALSRQRLRIILPIREYVDLRLDAFRTGL"
ORIGIN
Query Match 39.3%; Score 118.2; DB 9; Length 596;
Best Local Similarity 97.6%; Pred. No. 1.1e-13;
Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 18 TTGAGATTATTTTTCAGTCTGCTGGATACCTTTGTAGAAAACACGAAAGAACTTTCAA 77
Db 421 TTCAAAATATTTATTTTCTGCTGGAATACTTTGTAGAAAACCAATGAAGAACTTTCAA 480
OY 78 GCCTGGGAAGGCGTCGTAAGAAATTCAGTTCGTCTCCAGACAGCTTCGGCCATCCT 137
Db 481 GCCTGGGAAGGCGTCGTAAGAAATTCAGTTCGTCTCCAGACAGCTTCGGCCATCCT 540
OY 138 TTG 140
Db 541 TTG 543
RESULT 14
AF529818 596 bp mRNA linear PRI 17-SEP-2002
LOCUS AF529818
DEFINITION Homo sapiens clone Ramos 4 truncated AID (AID) mRNA, complete cds.
ACCESSION AF529818
VERSION AF529818.1 GI:22297222
KEYWORDS
SOURCE
Organism Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 596)
Martin,A. and Scharff,M.D.
Somatic hypermutation of the AID transgene in B and non-B cells
Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
2 (bases 1 to 596)
Martin,A. and Scharff,M.D.
Direct Submission
Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA
FEATURES
source
1. 596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 4"
1. .596
/gene="AID"
1. .570
/gene="AID"
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1

```

/product="truncated AID"
/protein_id="AA05405.1"
/db_xref="GI:22297224"
/translation="MDSLMMRRKFLYQFKNVMAKRRRTTYLCYVKKRDSATSFSL
DFGLRNKNGCHELLFLRYISDWLDPGCYRVTFWTSPCYDCAHVAADPLRNP
NLALRITFARLYFCEDEHKAPBEGRLRLHRAQVOIAIMTFDYFCYNTFVENHERTEK
AMEGLHNSVRLSPQLRILLPLYEVDL"

ORIGIN

Query Match 39.3%; Score 118.2; DB 9; Length 596;
Best Local Similarity 97.6%; Pred. No. 1.1e-13;
Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 TTGAAGATTATTTTCTGCTGGAATCTTTGTAGAAAACGAGAAAGACTTTCAAA 77
|||
DB 421 TTCAAGATTATTTTCTGCTGGAATCTTTGTAGAAAACGAGAAAGACTTTCAAA 480
|||
QY 78 GCTGGGAAGGGCTGCATGAAAATTTCAGTTCCTCTCCAGACAGCTTCGGCGCATCCTT 137
|||
DB 481 GCTGGGAAGGGCTGCATGAAAATTTCAGTTCCTCTCTCCAGACAGCTTCGGCGCATCCTT 540
|||
QY 138 TTG 140
|||
DB 541 TTG 543

RESULT 15

AF529819 596 bp mRNA linear PRI 17-SEP-2002
LOCUS Homo sapiens clone Ramos 5 AID (AID) mRNA, partial cds.

DEFINITION AF529819

ACCESSION AF529819

VERSION AF529819.1 GI:22297225

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..596

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="Ramos 5"

1..>596

/gene="AID"

1..>596

/note="Integrated into Burkitt's lymphoma cell line Ramos"

/product="AID"

/protein_id="AA05406.1"

/db_xref="GI:22297226"

/translation="MDSLMMRRKFLYQFKNVMAKRRRTTYLCYVKKRDSATSFSL
DFGLRNKNGCHELLFLRYISDWLDPGCYRVTFWTSPCYDCAHVAADPLRNP
NLALRITFARLYFCEDEHKAPBEGRLRLHRAQVOIAIMTFDYFCYNTFVENHERTEK
AMEGLHNSVRLSPQLRILLPLYEVDLDAFRTTGL"

ORIGIN

Query Match 39.3%; Score 118.2; DB 9; Length 596;

Best Local Similarity 97.6%; Pred. No. 1.1e-13;

Matches 120; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 TTGAAGATTATTTTCTGCTGGAATCTTTGTAGAAAACGAGAAAGACTTTCAAA 77

DB 421 TTCAAGATTATTTTCTGCTGGAATCTTTGTAGAAAACGAGAAAGACTTTCAAA 480
|||
QY 78 GCTGGGAAGGGCTGCATGAAAATTTCAGTTCCTCTCCAGACAGCTTCGGCGCATCCTT 137
|||
DB 481 GCTGGGAAGGGCTGCATGAAAATTTCAGTTCCTCTCTCCAGACAGCTTCGGCGCATCCTT 540
|||
QY 138 TTG 140
|||
DB 541 TTG 543

Search completed: March 12, 2004, 21:22:25
Job time : 949.629 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 13:19:34 ; Search time 958.46 Seconds
(without alignments)
9378.080 Million cell updates/sec

Title: US-09-966-880A-10_COPY_6264_6564

Perfect score: 301
Sequence: 1 gcataaaggagcactagta.....cttcacaaggcggtcaag 301

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthm:*
3: em_estim:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inv:*
19: em_ges_pln:*
20: em_ges_vrt:*
21: em_ges_fun:*
22: em_ges_mam:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_ror:*
26: em_ges_phg:*
27: em_ges_vrl:*
28: gb_ges1:*
29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 79.6 | 26.4 | 187 | AA249004 | AA249004 12585.seq |
| 2 | 75.2 | 25.0 | 476 | AO671876 | AO671876 HS_5462_A |
| 3 | 74.2 | 24.7 | 674 | AG051626 | AG051626 Pan trog1 |
| 4 | 74 | 24.6 | 894 | BQ103839 | BQ103839 ESTCENRDR2 |

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 5 | 71.2 | 23.7 | 630 | 14 | H59882 |
| 6 | 70.4 | 23.4 | 617 | 9 | AL038727 |
| 7 | 70.4 | 23.4 | 651 | 12 | BM983994 |
| 8 | 70.4 | 23.4 | 783 | 12 | BG741476 |
| 9 | 68.8 | 22.9 | 303 | 10 | AM948684 |
| 10 | 68.8 | 22.9 | 548 | 13 | EX362298 |
| 11 | 68.8 | 22.9 | 737 | 13 | EX094095 |
| 12 | 68.8 | 22.9 | 768 | 13 | CB311387 |
| 13 | 68.8 | 22.9 | 913 | 13 | EX374004 |
| 14 | 68.8 | 22.9 | 930 | 13 | BQ171790 |
| 15 | 67.6 | 22.5 | 699 | 28 | AC420050 |
| 16 | 67.6 | 22.5 | 446 | 10 | BF760170 |
| 17 | 66.8 | 22.2 | 360 | 28 | AO012711 |
| 18 | 66.8 | 22.2 | 461 | 28 | AO078850 |
| 19 | 66.8 | 22.2 | 1201 | 9 | AL573148 |
| 20 | 66.4 | 22.1 | 443 | 9 | AA292629 |
| 21 | 66.4 | 21.9 | 682 | 12 | BC538032 |
| 22 | 65.2 | 21.7 | 454 | 28 | AO322538 |
| 23 | 65.2 | 21.7 | 813 | 14 | CB312335 |
| 24 | 65.2 | 21.7 | 721 | 10 | AM954122 |
| 25 | 65 | 21.6 | 741 | 12 | BM674784 |
| 26 | 64.6 | 21.5 | 311 | 28 | B43977 |
| 27 | 64.6 | 21.5 | 440 | 28 | AO734239 |
| 28 | 64.2 | 21.3 | 434 | 28 | AO737126 |
| 29 | 64 | 21.3 | 893 | 13 | BU542546 |
| 30 | 63.8 | 21.2 | 339 | 14 | T87005 |
| 31 | 63.8 | 21.2 | 629 | 13 | EX111369 |
| 32 | 63.8 | 21.2 | 690 | 28 | AO028551 |
| 33 | 63 | 20.9 | 366 | 12 | BT789069 |
| 34 | 63 | 20.9 | 559 | 12 | BT962507 |
| 35 | 62 | 20.6 | 305 | 9 | AA385457 |
| 36 | 62 | 20.6 | 401 | 9 | AI252771 |
| 37 | 62 | 20.6 | 408 | 9 | AA281997 |
| 38 | 62 | 20.6 | 452 | 9 | AI379101 |
| 39 | 62 | 20.6 | 509 | 9 | AI827559 |
| 40 | 62 | 20.6 | 913 | 13 | BQ883450 |
| 41 | 61.8 | 20.5 | 674 | 29 | AG182249 |
| 42 | 61.4 | 20.4 | 346 | 28 | AO012867 |
| 43 | 61.2 | 20.3 | 421 | 28 | AO016082 |
| 44 | 61.2 | 20.3 | 608 | 29 | AG161212 |
| 45 | 61.2 | 20.3 | 861 | 14 | CD516515 |

ALIGNMENTS

RESULT 1
AA249004/c
LOCUS
DEFINITION
12585.seq.F Human fetal heart, lambda ZAP Express Homo sapiens cDNA 5', mRNA sequence.
ACCESSION
AA249004
VERSION
AA249004.1 GI:1880979
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
Llew.C.C.
CDNAs from human fetal heart (1997)
Unpublished (1997)
CONTACT: Llew CC
Brigham and Women's Hospital
Harvard Medical School
75 Francis St. Boston, MA 02115, USA
Tel: 617328915
Fax: 6179750995
Email: cliew@rics.bwh.harvard.edu
PCR Primers
FORWARD: 5' GCCAGCTGGAATTAACCTCACTAAAGG 3'
REVERSE: 5' CCGTGAATTTGATACGACTCACTAAGGCGG 3'
Seq primer: 5' GAAATTAACCTCACTAAAGG 3'

Query Match 24.7%; Score 74.2; DB 29; Length 674;
 Best Local Similarity 71.9%; Pred. No. 2e-09;
 Matches 97; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 87 TGTGATATATGATATATATATAGTTTGTTCATAGTCTGCTGATTAACCCCTC 146
 DB 266 TGTGATATATGATATATATATAGTTTGTTCATAGTCTGCTGATTAACCCCTC 325

QY 147 ACCCTGTTACATCATTTGTTTAAGGTGATGTTAGGCTCAGAGAAACTCT 206
 DB 326 GCCCTGTTACATCATTTGTTTAAGGTGATGTTAGGCTCAGAGAAACTCT 385

QY 207 CTCTCTACCTCTC 221
 DB 386 GACCTTCTCTGCCC 400

RESULT 4
 LOCUS B0103839 894 bp mRNA linear EST 20-JAN-2003
 DEFINITION ESTCRD227 Human cervical cDNA (differential display PCR) Homo sapiens cDNA 3', mRNA sequence.
 ACCESSION B0103839
 VERSION B0103839.1 GI:27804332
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 894)
 Vernon,S.D. and Unger,E.R.
 Gene expression profiling of keratinocyte differentiation by fluorescent differential display PCR
 Unpublished (2003)
 JOURNAL Contact: Rajeevan MS
 Division of Viral and Rickettsial diseases
 Centers for Disease Control and Prevention
 1600 Clifton Rd NE, Mail stop G-41, Atlanta, GA 30333, USA
 Email: moradocdc.gov
 A transcript up-regulated in differentiating compared to proliferating keratinocytes. No significant match identified by protein domain search
 Insert Length: 894 Std Error: 0.00
 Seq primer: M13 Forward
 High quality sequence stop: 894
 POLVA=Yes.

FEATURES
 source Location/Qualifiers
 1..894
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_type="cervical keratinocytes"
 /cell_line="Subclone 20861 of W12 cell line"
 /clone_1ib="Human cervical cDNA (differential display PCR)"
 /note="Genes differentially expressed during keratinocyte differentiation were identified by comparing DD-PCR profiles of cells grown under non-differentiating (monolayers) and differentiating (raft cultures) conditions. Differentially expressed bands were excised and sequenced for functional annotation"

ORIGIN
 Query Match 24.6%; Score 74; DB 13; Length 894;
 Best Local Similarity 75.4%; Pred. No. 2.3e-09;
 Matches 92; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 78 GATAGGGTGTGATATATATATATAGTTTGTTCATAGTCTGCTGATTA 137
 DB 272 GACATGGGTAAGTGTGATATATATATAGTTTGTTCATAGTCTGCTGATTA 331

QY 138 AACCCCTCACCTTGTACAGTCATTTGTTATATAGGTGATGTTAGCCTCAGAG 197
 DB 332 ACTCCCATAGCCCTTATTATAGTGTGTTTATATATGTTGGTGTATACCTCAGTGG 391

QY 198 CA 199
 DB 392 CA 393

RESULT 5
 H59882
 LOCUS H59882 630 bp mRNA linear EST 06-OCT-1995
 DEFINITION Y10409.r1 Soares fetal liver spleen INFILS Homo sapiens cDNA clone IMAGE:204281 5' similar to gb|MF7919|HMANE53 Human carcinoma cell-derived Alu RNA transcript. (rRNA); gb:X52075_rna3 LEUKOSIALIN PRECURSOR (HUMAN);, mRNA sequence.
 ACCESSION H59882
 VERSION H59882.1 GI:1012714
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 630)
 Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Maitra,M., Parsons,U., Rikkin,U., Ronfling,T., Soares,M., Tan,F., Tsvetakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
 The Mashu-Merck EST Project
 Unpublished (1995)
 JOURNAL Contact: Wilson RK
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 851
 High quality sequence stops: 461
 Source: IMAGE Consortium, LNL
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 851 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 461.

FEATURES
 source Location/Qualifiers
 1..630
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:3773412"
 /db_xref="taxon:9606"
 /clone="IMAGE:204281"
 /sex="male"
 /dex_stage="20 week-post conception fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_1ib="Soares fetal liver spleen INFILS"
 /note="Organ: Liver and Spleen; Vector: pTR73D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' ACTCGAAGATATATATATAGTCTTCTTTTCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pTR73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN
 Query Match 23.7%; Score 71.2; DB 14; Length 630;
 Best Local Similarity 73.4%; Pred. No. 1.3e-08;
 Matches 91; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 76 CTGATTAAGGTTGTGATATATATATATATAGTTTGTTCATAGTCTGCTGATTA 135

DB 179 CAGATGGAGGCTGATACGATCAGATGATATGTTTCATCCGCAATTCCTGGCTCA 238

QY 136 TAAACCCCTCACCCTTGTACGATTTGTTATAGGTTAGGCTTACGA 195

DB 229 TAATCCCATAGCCCTTATACATGTTTGTATCATGTTGGGAGTAGTGGCTCAGG 298

QY 196 AGCA 199

DB 299 AGCA 302

RESULT 5

LOCUS AL038727 617 bp mRNA linear EST 04-SEP-2003

DEFINITION DKFZP566L1846_r1 566 (synonym: hfk2) Homo sapiens cDNA clone

ACCESSION AL038727

VERSION AL038727.1 GI:5407877

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 617)

AUTHORS Oetemaeder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and Wiemann, S.

TITLE EST (Oetemaeder, et al.)

JOURNAL Unpublished (1999)

COMMENT Contact: MIPS

FEATURES

source

1. 617

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFZP566L1846"

/tissue_type="Kidney"

/dev_stage="fetal"

/lab_host="XI-2blue"

/clone_1ib="566 (synonym: hfk2)"

/note="Vector: pAMP1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 23.4%; Score 70.4; DB 9; Length 617;

Best Local Similarity 83.6%; Pred. No. 2.1e-08;

Matches 92; Conservative 0; Mismatches 16; Indels 2; Gaps 1;

QY 188 GCTCAGAGCAAACTCTCTCTACCTTCCAGCCCTCTCTCTCTGACCTCAT 247

DB 342 GCCCAGAGCAAAAGTTCTCTCTGACCTTCTCTGCTCTCTCTCTCTCTTTCAT 401

QY 248 TCTTCCCTGAGGC--CACATGAAACATGAAATCTCTTCCACAGGCGG 295

DB 402 TCTCCCAAGGCTACCCATGAAACTGAAATCCTCTTCCCAAGGCGG 451

RESULT 7

LOCUS BM983994/c 651 bp mRNA linear EST 20-FEB-2003

DEFINITION UI-CF-DUI-aba-1-17-0-UI.s1 UI-CF-DUI Homo sapiens cDNA clone

ACCESSION UI-CF-DUI-aba-1-17-0-UI 3, mRNA sequence.

BM983994

VERSION BM983994.1 GI:19609061

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 651)

AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 9704447

PUBMED 8889548

COMMENT Contact: McCray, PB

McCray Lab

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welch, University of Iowa

CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com).

The following repetitive elements were found in this cDNA sequence: 23-564, >MER39B

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES

source

Location/Qualifiers

1. 651

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-CF-DUI-aba-1-17-0-UI"

/tissue_type="Primary Lung Epithelial Cells"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_1ib="UI-CF-DUI"

/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-DUI is a normalized cDNA library containing the following tissue(s): Primary Lung Epithelial Cells The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCGTGTAGGC.

TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368

TAG_LIB=UI-CF-DUI

TAG_SEQ=GCGTGTAGGC"

ORIGIN

Query Match 23.4%; Score 70.4; DB 12; Length 651;

Best Local Similarity 83.6%; Pred. No. 2.1e-08;

Matches 92; Conservative 0; Mismatches 16; Indels 2; Gaps 1;

QY 188 GCTCAGAGCAAACTCTCTCTACCTTCCAGCCCTCTCTCTGACCTCAT 247

DB 538 GCCCAGAGCAAAAGTTCTCTCTGACCTTCTCTGCTCTCTCTCTCTTTCAT 479

QY 248 TCTTCCCTGAGGC--CACATGAAACATGAAATCTCTTCCACAGGCGG 295

DB 478 TCTCCCAAGGCTACCCATGAAACTGAAATCCTCTTCCCAAGGCGG 429

RESULT 8

BG741476 783 bp mRNA linear EST 15-MAY-2001
 LOCUS BG741476
 DEFINITION 602632271F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4777553 5',
 mRNA sequence.
 ACCESSION BG741476
 VERSION BG741476.1 GI:14052129
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgs.nci.nih.gov/
 1 (bases 1 to 783)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: rgs@nih.gov
 Tissue Procurement: James Cleaver, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA
 Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM10631 row: n column: 18
 High quality sequence stop: 760.
 Location/Qualifiers
 1. 783
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4777553"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI CGAP Skn3"
 /note="Organ: skin; Vector: pCMV-SPORT6, Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.5kb. Library constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."
 ORIGIN
 Query Match 23.4%; Score 70.4; DB 12; Length 783;
 Best Local Similarity 83.6%; Pred. No. 2.2e-08;
 Matches 92; Conservative 0; Mismatches 16; Indels 2; Gaps 1;
 QY 188 GCTCAGAGCAAACTCTCTCTCCTCCACCTTCCAGCCCTCGTCTGCAACCTCAT 247
 DB 243 GCCCAGAGCAAAAGTTCTCTCTGACCTTCTCTGCTGCTGCTGCTGCTTTCAT 302
 QY 248 TCTTCCCTGAGGC--CACTAGAACTAGATCTCTTCCACAGGCG 295
 DB 303 TCTCCCCAGAGCTACCCATAGAACTAGATCCCTTCCCCAAGGCG 352
 RESULT 9
 AM948684 303 bp mRNA linear EST 31-MAY-2000
 LOCUS AM948684/c
 DEFINITION QV0-FT0002-050500-230-f04 FT0002 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM948684
 VERSION AM948684.1 GI:8126458
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 303)
 Nias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R.,
 Dias, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 PUBMED 10737800
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?file=et2-qv0-FT0002-050
 500-230-f04&3=2000-05-05&4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 303.
 Location/Qualifiers
 1. 303
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="FT0002"
 /note="Organ: prostate tumor; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORFESTS PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 ORIGIN
 Query Match 22.9%; Score 68.8; DB 10; Length 303;
 Best Local Similarity 82.7%; Pred. No. 5.2e-08;
 Matches 91; Conservative 0; Mismatches 17; Indels 2; Gaps 1;
 QY 188 GCTCAGAGCAAACTCTCTCTCCTCCACCTTCCAGCCCTCGTCTGCAACCTCAT 247
 DB 234 GCCCAGAGCAAAAGTTCTCTCTGACCTTCTCTGCTGCTGCTGCTGCTTTCAT 175
 QY 248 TCTTCCCTGAGGC--CACTAGAACTAGATCTCTTCCACAGGCG 295
 DB 174 TCTCCCCAGAGCTACCCATAGAACTAGATCCCTTCCCCAAGGCG 125
 RESULT 10
 BX362298 548 bp mRNA linear EST 05-MAY-2003
 LOCUS BX362298
 DEFINITION BX362298 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
 ACCESSION BX362298
 VERSION BX362298.1 GI:30368597
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 548)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5793.f for
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DJ013AA02Pc1cluster=5793.f. Contact :
 Feng Liang Email : fliang@life-tech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

QY 188 GCCTCAGAGCAAACTCTCTCTCTCCTCAGCCCTCCTCTCTGACCTCAT 247
 DB 295 GCGCCGAGAGCAAAATTTCTCTCTCCTCTCCTGCGCTTGTCTTCTGACCTCAT 354
 QY 248 TCTTCCCTGAGGC--CACATAGAACTAGATCTCTCTTCCACAGGCGG 295
 DB 355 TCTCCCGCAAGGCTACCATAGAACTAGATCTCTTCCCGCAAGGAG 404

RESULT 13
 BX374004 913 bp mRNA linear EST 08-MAY-2003
 LOCUS BX374004 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
 DEFINITION Homo sapiens cDNA clone CS0D013YA03 3-PRIME, mRNA sequence.
 ACCESSION BX374004
 VERSION BX374004.1 GI:30450260
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 913)
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5793.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0BAK070CG09NML&cluster=5793.f. Contact :
 Feng Liang Email: fliang@life.com/Invitrogen Corporation 1600
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID: CS0BAK070CG09NML.
 Location/Qualifiers
 1..913
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D013YA03"
 /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /cell_line="JURKAT"
 /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
 source

ORIGIN
 Query Match 22.9%; Score 68.8; DB 13; Length 913;
 Best Local Similarity 82.7%; Pred. No. 6e-08;
 Matches 91; Conservative 0; Mismatches 17; Indels 2; Gaps 1;

QY 188 GCCTCAGAGCAAACTCTCTCTCTCCTCAGCCCTCCTCTCTGACCTCAT 247
 DB 520 GCGCCGAGAGCAAAATTTCTCTCTCCTCTCCTGCGCTTGTCTTCTGACCTCAT 461
 QY 248 TCTTCCCTGAGGC--CACATAGAACTAGATCTCTCTTCCACAGGCGG 295
 DB 460 TCTCCCGCAAGGCTACCATAGAACTAGATCTCTTCCCGCAAGGAG 411

RESULT 14
 BQ717190 930 bp mRNA linear EST 16-JUL-2002
 LOCUS BQ717190
 DEFINITION AGNCOURT 8232763 Lupski sympathetic_trunk Homo sapiens cDNA clone
 IMAGE:6191636 5', mRNA sequence.
 ACCESSION BQ717190
 VERSION BQ717190.1 GI:21856087

KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 930)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LHAM1352 row: f column: 21
 High quality sequence stop: 625.
 Location/Qualifiers
 1..930
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6191636"
 /sex="male"
 /tissue_type="sympathetic trunk"
 /dev_stage="adult", 16 yr"
 /lab_host="DH10B"
 /clone_lib="Lupski sympathetic trunk"
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
 NotI; Site 2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCAGCGCTCG-3' and
 5'-GACATGTTCTGATCGGAGCGCGGCTT(15)-3'. Size selected >
 1 kb for average insert length 1.9 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine); available through Life
 Technologies."

ORIGIN
 Query Match 22.9%; Score 68.8; DB 13; Length 930;
 Best Local Similarity 82.7%; Pred. No. 6e-08;
 Matches 91; Conservative 0; Mismatches 17; Indels 2; Gaps 1;

QY 188 GCCTCAGAGCAAACTCTCTCTCTCCTCAGCCCTCCTCTCTGACCTCAT 247
 DB 503 GCGCCGAGAGCAAAATTTCTCTCTCCTCTCCTGCGCTTGTCTTCTGACCTCAT 562
 QY 248 TCTTCCCTGAGGC--CACATAGAACTAGATCTCTCTTCCACAGGCGG 295
 DB 563 TCTCCCGCAAGGCTACCATAGAACTAGATCTCTTCCCGCAAGGAG 612

RESULT 15
 AQ420050 699 bp DNA linear GSS 23-MAR-1999
 LOCUS AQ420050
 DEFINITION RPCT-11-205H11.TV RPCT-11 Homo sapiens genomic clone
 RPCT-11-205H11, genomic survey sequence.
 ACCESSION AQ420050
 VERSION AQ420050.1 GI:4477774
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 699)
 Zhao, S., Adams, M.D., Niernan, W., Malek, J., de Jong, P. and
 Venter, J.C.
 Use of BAC End Sequences from Library RPCT-11 for Sequence-Ready
 Map Building

JOURNAL
COMMENT
Unpublished (1997)
Other GSSs: RPCI-11-205H11.TV
Contact: Shaying Zhao, William Niernman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.
Location/Qualifiers

1. .699
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7578514"
/db_xref="taxon:9606"
/clone="RPCI-11-205H11"
/sex="Male"
/cell_type="lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"

ORIGIN

Query Match 22.5%; Score 67.8; DB 28; Length 699;
Best Local Similarity 73.0%; Pred. No. 1.1e-07;
Matches 119; Conservative 0; Mismatches 32; Indels 12; Gaps 2;
QY 87 TGTGATATATGATATATATAGTTTGTTCATAGTTGCTTATTAACCCCTC 146
DB 441 TATGATTTTATGATACATGTTTTCATCCACGCTCTGAGCTCATTAATCTCTATA 382
QY 147 ACCCTGTACAGTCATTTGTATAAGTT-GGATGTTTAGGCTCAGAGCAAAACTC 205
DB 381 ACCCTGTATTTGCTTTTGTATAAGTTGGGCACTTAGGCTCAGAGGAGGCTC 322
QY 206 -----TCTCTCTACCTTTCTCAGCCCTCTGTCT 237
DB 321 AGGAGCAGATCTCTGACCTTCTCGCTTCTTACTT 279

Search completed: March 13, 2004, 00:38:48
Job time : 962.585 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 12:27:59 ; Search time 947.629 Seconds

(without alignments)
13767.261 Million cell updates/sec

Title: US-09-966-880A-10_COPY_6264_6564

Perfect score: 301
Sequence: 1 gacccaagagacacttagta.....ctccacaagcgctcaag 301Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
GenEmbl:
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sta:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: gb_wa:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_cm:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sta:*
28: em_un:*
29: em_vl:*
30: em_hg_hum:*
31: em_hg_inh:*
32: em_hg_other:*
33: em_hg_mus:*
34: em_hg_pin:*
35: em_hg_rod:*
36: em_hg_mam:*
37: em_hg_vrt:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|---------------|--------------------|
| 1 | 301 | 100.0 | 6564 | 6 BD016835 | BD016835 Novel cyt |
| 2 | 301 | 100.0 | 71132 | 9 AC092184 | AC092184 Homo sapi |
| 3 | 89.4 | 29.7 | 39395 | 9 AC004212 | AC004212 Homo sapi |
| 4 | 89.4 | 29.7 | 76225 | 9 AL773551 | AL773551 Human DNA |
| 5 | 89.4 | 29.7 | 112018 | 9 AB023050 | AB023050 Homo sapi |
| 6 | 89.4 | 29.7 | 171627 | 9 AL662797 | AL662797 Human DNA |
| 7 | 89.4 | 29.7 | 156617 | 9 AL662848 | AL662848 Human DNA |
| 8 | 89.4 | 29.7 | 200000 | 9 AP000512 | AP000512 Homo sapi |
| 9 | 87.8 | 29.2 | 300050 | 9 AB100084 | AB100084 Pan trogl |
| 10 | 83.8 | 27.8 | 481 | 9 HSDJ93K2 | HSDJ93K2 |
| 11 | 83.8 | 27.8 | 113804 | 9 AC025610 | AC025610 Homo sapi |
| 12 | 83.8 | 27.8 | 170877 | 2 AC021655 | AC021655 Homo sapi |
| 13 | 82.6 | 27.4 | 143133 | 2 AC092911 | AC092911 Homo sapi |
| 14 | 82.6 | 27.4 | 163862 | 2 AC026559 | AC026559 Homo sapi |
| 15 | 82.6 | 27.4 | 165082 | 9 AC074044 | AC074044 Homo sapi |
| 16 | 82.6 | 27.4 | 194006 | 9 AC024108 | AC024108 Homo sapi |
| 17 | 82.6 | 27.4 | 229396 | 9 AC069431 | AC069431 Homo sapi |
| 18 | 81.6 | 27.1 | 123742 | 9 AC036127 | AC036127 Homo sapi |
| 19 | 81.6 | 27.1 | 152433 | 2 BX890549 | BX890549 Dario rer |
| 20 | 81.2 | 27.0 | 123057 | 9 AL353741 | AL353741 Human DNA |
| 21 | 81.2 | 27.0 | 199517 | 2 AC131265 | AC131265 Homo sapi |
| 22 | 80.6 | 26.8 | 59561 | 2 AC021316 | AC021316 Homo sapi |
| 23 | 80.6 | 26.8 | 173535 | 9 AC104759 | AC104759 Homo sapi |
| 24 | 80.6 | 26.3 | 173986 | 9 AC006236 | AC006236 Homo sapi |
| 25 | 79.2 | 26.1 | 176368 | 2 AC140859 | AC140859 Homo sapi |
| 26 | 78.6 | 26.1 | 185737 | 2 AC145616 | AC145616 Homo sapi |
| 27 | 78.6 | 26.1 | 197637 | 9 AC093141 | AC093141 Homo sapi |
| 28 | 78.6 | 26.1 | 207614 | 9 AC098800 | AC098800 Homo sapi |
| 29 | 78.4 | 26.0 | 171683 | 9 AC105179 | AC105179 Homo sapi |
| 30 | 77.8 | 25.8 | 140137 | 2 AC027611 | AC027611 Homo sapi |
| 31 | 77.8 | 25.8 | 151847 | 2 AC022792 | AC022792 Homo sapi |
| 32 | 77.8 | 25.8 | 165378 | 9 AC129805 | AC129805 Homo sapi |
| 33 | 77.6 | 25.8 | 165378 | 9 AP005360 | AP005360 Homo sapi |
| 34 | 77.6 | 25.8 | 152440 | 9 AC104233 | AC104233 Homo sapi |
| 35 | 77.6 | 25.8 | 162241 | 9 AC022909 | AC022909 Homo sapi |
| 36 | 77.6 | 25.8 | 184533 | 9 AC093592 | AC093592 Homo sapi |
| 37 | 77.6 | 25.8 | 192573 | 9 AC073344 | AC073344 Homo sapi |
| 38 | 77.4 | 25.7 | 158562 | 9 AL603888 | AL603888 Human DNA |
| 39 | 77.4 | 25.5 | 110000 | 2 AL929091_02 | Continuation (3 of |
| 40 | 76.8 | 25.5 | 110000 | 2 AL929091_03 | Continuation (4 of |
| 41 | 76.8 | 25.5 | 155889 | 2 AC025642 | AC025642 Homo sapi |
| 42 | 76.8 | 25.4 | 162347 | 9 AL356108 | AL356108 Human DNA |
| 43 | 76.4 | 25.4 | 216451 | 9 AC098818 | AC098818 Homo sapi |
| 44 | 76.4 | 25.4 | 219207 | 2 AC009685 | AC009685 Homo sapi |
| 45 | 76.4 | 25.4 | 219207 | 2 AC009685 | AC009685 Homo sapi |

ALIGNMENTS

```
RESULT 1
BD016835
LOCUS      BD016835
DEFINITION Novel cytidine deaminase.
ACCESSION  BD016835
VERSION    BD016835.1 GI:22558011
KEYWORDS   Novel cytidine deaminase
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 6564)
AUTHORS   Honjo, T. and Muramatsu, M.
TITLE     Novel cytidine deaminase
JOURNAL   Patent: JP 2001245669-A/8 11-SEP-2001;
```

COMMENT JAPAN TOBACCO INC. TASUKU HONJO
OS Homo sapiens (human)
FN JP 2001245669-A/8
PD 11-SEP-2001
PF 28-MAR-2000 JP 2000092981
PI TASUKU HONJO MASAMICHI MIMAMATSU
PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12,
PC A61P17/00,
PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12N9/78, C12P21/02, C12P21/08 // (C12N1/21, C12R1/19), PC
(C12N5/10, C12R1/91), C12N15/00, C12N5/00, (C12N5/00, C12R1/91) CC
FH Key Location/Qualifiers.
FEATURES
source 1. .6564
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 301; DB 6; Length 6564;
Best Local Similarity 100.0%; Pred. No. 1e-71;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCATCAAGAGCACTTAGTAAAGTGAAGAACCAAGCAAGAGTGGAGAAACAC 60
DB 6264 GCATCAAGAGCACTTAGTAAAGTGAAGAACCAAGCAAGAGTGGAGAAACAC 6323
QY 61 TTGCAANTCATATATCTGATTAAGGTTGGATTTATGATATATATATGTTTTC 120
DB 6324 TTGCAANTCATATATCTGATTAAGGTTGGATTTATGATATATATGTTTTC 6383
QY 121 ATAGTTCCTGGCTTATTAACCCCTCACCTTGTACAGTTCATTGTTAAGTTGAT 180
DB 6384 ATAGTTCCTGGCTTATTAACCCCTCACCTTGTACAGTTCATTGTTAAGTTGAT 6443
QY 181 GGTATAGGCTCAGAAACAACCTCTCTCTCACCCTTCCAGCCCTCTCTCTGAC 240
DB 6444 GGTATAGGCTCAGAAACAACCTCTCTCTCACCCTTCCAGCCCTCTCTCTGAC 6503
QY 241 ACCCTATTTCTTCCCTGAGGCGCATAGAACTGAACTCTCTCTTCCACAGGCGGTCAA 300
DB 6504 ACCCTATTTCTTCCCTGAGGCGCATAGAACTGAACTCTCTCTTCCACAGGCGGTCAA 6563
QY 301 G 301
DB 6564 G 6564

RESULT 2
LOCUS AC092184 71132 bp DNA linear PRI 12-UTN-2002
DEFINITION Homo sapiens 12 BAC RP11-438L7 (Roswell Park Cancer Institute Human
AC092184 AC013443
ACCESSION AC092184 complete sequence.
VERSION AC092184.7 GI:21206067
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 71132)
Muzny, D.M., Adams, C., Adio-Oduola, B., All-omen, F.R., Allen, C.,
Alsbrooks, S.L., Amaral, H.C., Are, J.R., Ayala, M., Banks, T.,
Barbier, J., Benton, J., Bimaga, K., Blankenburg, K., Bonin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhury, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Doutwalte, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,

REFERENCE
AUTHORS
TITLE
JOURNAL
Unpublished
2 (bases 1 to 71132)
Worley, K.C.
Direct Submission
Submitted (18-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 71132)
Worley, K.C.
Direct Submission
Submitted (18-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 71132)
Worley, K.C.
Direct Submission
Submitted (25-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 71132)
Worley, K.C.
Direct Submission
Submitted (12-UTN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 25, 2002 this sequence version replaced gi:20901754.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
STS are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,

unpublished.) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarly (expect $< 1e-34$) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES. This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found on the URL:

http://gc.bcm.tmc.edu/quality_info/genbank_annotation.html.

QUALSTAT-REPORT

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. .71132 |

```

repeat_region complement(6711. .6817)
/rpt_family="MIR"
repeat_region complement(7033. .7327)
/rpt_family="AluBx"
repeat_region 7607. .7745
/rpt_family="FLAM_A"
repeat_region 7980. .8047
/rpt_family="MIR"
repeat_region complement(8050. .8549)
/rpt_family="LTR47A"
repeat_region 8567. .8698
/rpt_family="MIR"
repeat_region complement(8815. .9101)
/rpt_family="AluSg"
repeat_region complement(11227. .12175)
/rpt_family="LTR5"
repeat_region 12625. .12900
/standard_name="87432"
repeat_region complement(13547. .13664)
/rpt_family="L1MB5"
repeat_region complement(13665. .13979)
/rpt_family="AluY"
repeat_region complement(13980. .14193)
/rpt_family="L1MB5"
repeat_region 14622. .14924
/rpt_family="AluY"
repeat_region complement(15813. .17136)
/rpt_family="L1Pa8"
repeat_region complement(17272. .17395)
/rpt_family="FLAM_C"
repeat_region 17426. .17575
/rpt_family="MIR"

```

| | | | | |
|-------------|---------|------------|-------|---------------|
| Query Match | 100.0%; | Score 301; | DB 9; | Length 71132; |
|-------------|---------|------------|-------|---------------|

Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-------|--|-------|
| Qy | 1 | GCATCAAAAGCACTTAGTAAAGAACTGAAAAAGACACACCACAGAAAGTGGAGAAAAACAC | 60 |
| Db | 46022 | GCATCAAAAGCACCTTAGTAAAGAAAGTAAAAAGACACACCACAGAAAGTGGAGAAAAACAC | 46081 |
| Qy | 61 | TTGCAAATCATATATCTGATTAAGGTTGTGATATTATGATATATATATAGTTTGTGCC | 120 |
| Db | 46082 | TTGCAAATCATATATCTGATTAAGGTTGTGATATTATGATATATATATAGTTTGTGCC | 46141 |
| Qy | 121 | ATAGTCTCGCTTAATPAAACCCCTCACCCCTGTTCACATCAATTGTATTAAGTTGAT | 180 |
| Db | 46142 | ATAGTCTCGCTTAATPAAACCCCTCACCCCTGTTCACATCAATTGTATTAAGTTGAT | 46201 |
| Qy | 181 | GGTTTAGGCTTCAGAAACAAACTCTCTCTCTCACTTTCAGACCCCTCTGTCTTGCC | 240 |
| Db | 46202 | GGTTTAGGCTTCAGAAACAAACTCTCTCTCTCACTTTCAGACCCCTCTGTCTTGCC | 46261 |
| Qy | 241 | ACCTCATTTCTCCCTTAGGCGCATATGAACATAGATCTCTTCCACAGGGGCTCAAA | 300 |
| Db | 46262 | ACCTCATTTCTCCCTTAGGCGCATATGAACATAGATCTCTTCCACAGGGGCTCAAA | 46321 |
| Qy | 301 | G 301 | |
| Db | 46322 | G 46322 | |

RESULT 3
AC004212
LOCUS
DEFINITION Homo sapiens clone UMGc:Y67c126 from 6p21, complete sequence.
ACCESSION AC004212
VERSION
KEYWORDS
SOURCE HTG
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 39395)
 AUTHORS Janer,M.M., Guillaudoux,T., Vu,Q., Kutyavin,T., Harter,H. and
 TITLE Large scale sequence analysis of the human MHC class I region
 JOURNAL Unpublished (1998)
 REMARK Fred Hutchinson Cancer Research Center
 The Clinical Research Division
 1100 Fairview Ave. N., P.O. Box 19024
 Seattle, WA 98109-1024

REFERENCE 2 (bases 1 to 39395)
 AUTHORS Geraghty,D.E. and Olson,M.V.
 TITLE Direct Submission
 JOURNAL Submitted (23-FEB-1998) Human Genome Center, University of
 Washington, Box 352145, Seattle, WA 98195, USA

REFERENCE 3 (bases 1 to 39395)
 AUTHORS Geraghty,D.E. and Olson,M.V.
 TITLE Direct Submission
 JOURNAL Submitted (17-JUL-1998) Human Genome Center, University of
 Washington, Box 352145, Seattle, WA 98195, USA

REMARK CONTACT: Daniel E. Geraghty (geraghty@fhrc.org)
 On Jul 17, 1998 this sequence version replaced gi:2905848.
 Overlapping sequences:
 5': UMGC:Y67C092
 3': UMGC:Y54C283

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

Double stranded (DS) coverage: 89.8%
 DS or two chemistry coverage: 100.0%
 Single stranded regions: 0

Sequence Validation:

This sequence has been validated by Multiple Complete Digest Mapping. Comparison of the experimentally derived map digest fragments with sequence-predicted fragments is given below. Small fragments below a variable cutoff (approximately 400-600bp) are not mapped and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragment groups are separated by dashed lines.

| Map | Seq | Map | Seq | Map | Seq |
|----------|----------|----------|----------|----------|----------|
| 13333.40 | 12934.00 | 4775.80 | 4608.00 | 5063.92 | 4866.00 |
| 10170.57 | 9750.00 | 8191.44 | 7867.00 | 560.75 | 556.00 |
| 3724.57 | 3591.00 | 11693.00 | 11114.00 | 2711.00 | 2644.00 |
| 5055.83 | 4873.00 | | | 13108.62 | 12618.00 |
| 3403.86 | 3333.00 | | | 876.14 | 866.00 |
| 1302.20 | 1294.00 | | | 10536.25 | 10080.00 |

FEATURES

source

1. 39395
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="6"
 /map="6p21"
 /clone="CGM1:A146A3"
 /sub_clone="UMGC:Y67C126"
 /cell_line="CGM1"

repeat_region /clone_lib="Wash U YAC library"
 repeat_region complement(11..498)
 repeat_region /rpt_family="Alu"
 repeat_region complement(11793..2035)
 repeat_region /rpt_family="Alu"
 repeat_region complement(2072..2364)
 repeat_region /rpt_family="Alu"
 repeat_region 2902..3188
 repeat_region /rpt_family="Alu"
 repeat_region complement(3297..3710)
 repeat_region /rpt_family="MLT2B2"
 repeat_region complement(3872..4138)
 repeat_region /rpt_family="Alu"
 repeat_region complement(5128..5284)
 repeat_region /rpt_family="Alu"
 repeat_region complement(9979..10294)
 repeat_region /rpt_family="Alu"
 repeat_region 10379..10492
 repeat_region /rpt_family="MIR"
 repeat_region 10872..11146
 repeat_region /rpt_family="MIR"
 repeat_region complement(11926..12095)
 repeat_region /rpt_family="Alu"
 repeat_region complement(12242..12529)
 repeat_region /rpt_family="Alu"
 repeat_region 12823..12874
 repeat_region /rpt_family="Alu"
 repeat_region complement(13004..13295)
 repeat_region /rpt_family="Alu"
 repeat_region 14863..15437
 repeat_region /rpt_family="Alu"
 repeat_region 16992..17282
 repeat_region /rpt_family="Alu"
 repeat_region complement(17924..18229)
 repeat_region /rpt_family="Alu"
 repeat_region 18513..18732
 repeat_region /rpt_family="Alu"
 repeat_region 19271..19498
 repeat_region /rpt_family="Alu"
 repeat_region 20131..21084
 repeat_region /rpt_family="LTR12"
 repeat_region 21145..21195
 repeat_region /rpt_family="Alu"
 repeat_region 21196..21470
 repeat_region /rpt_family="Alu"
 repeat_region 22542..22722
 repeat_region /rpt_family="Alu"
 repeat_region 23189..23488
 repeat_region /rpt_family="Alu"
 repeat_region 23542..23631
 repeat_region /rpt_family="Alu"
 repeat_region complement(24444..25026)
 repeat_region /rpt_family="Alu"
 repeat_region complement(26092..26389)
 repeat_region /rpt_family="Alu"
 repeat_region complement(26824..27046)
 repeat_region /rpt_family="Alu"
 repeat_region complement(27109..27405)
 repeat_region /rpt_family="Alu"
 repeat_region 27692..27976
 repeat_region /rpt_family="Alu"
 repeat_region 28990..29615
 repeat_region /rpt_family="LTR12"
 repeat_region complement(29635..29824)
 repeat_region /rpt_family="MER39"
 repeat_region complement(29635..29751)
 repeat_region /rpt_family="MER21"
 repeat_region complement(29853..30134)
 repeat_region /rpt_family="Alu"
 repeat_region complement(30162..30221)
 repeat_region /rpt_family="MER39"
 repeat_region 30431..30708
 repeat_region /rpt_family="Alu"

repeat_region complement(30836, .31706)
/rpt_family="MER21"
repeat_region complement(30895, .31195)
/rpt_family="MER39"
repeat_region complement(32270, .32547)
/rpt_family="Alu"
repeat_region complement(32599, .33195)
/rpt_family="Alu"
repeat_region complement(33661, .33946)
/rpt_family="Alu"
repeat_region 34848, .35115
/rpt_family="Alu"
repeat_region 35730, .35916
/rpt_family="Alu"
repeat_region 36293, .36408
/rpt_family="Alu"
repeat_region 37656, .37902
/rpt_family="MER42"
variation 38541
/note="high quality variation versus 5' overlapping clone"
repeat_region 38547, .38701
/rpt_family="Alu"
repeat_region 38843, .39083
/rpt_family="Alu"
variation 39213
/note="high quality variation versus 3' overlapping clone"

ORIGIN

Query Match 29.7%; Score 89.4; DB 9; Length 39395;
Best Local Similarity 73.7%; Pred. No. 6.9e-14;
Matches 132; Conservative 0; Mismatches 36; Indels 11; Gaps 1;

Qy 87 TGTGATATATGATATATATAGCTTTTGTCCATGTTCTCGGCTTATAAACCCTC 146
Db 30811 TGTGCTATATATATATATAGCTTTTGTCCATGTTCTCGGCTTATAAACCCTC 30870
Qy 147 ACCCTGTTACAGCATTTGTTATAGGTTGAGTGGCTTCAAGACAAACTC- 205
Db 30871 GCCCTGTTACCATCTGTTGTTATAGTGGCTTCAAGACCAAGAGCTCA 30930
Qy 206 -----TCTCTCAACCTTCTCCAGCCCTCTGCTCTGCAACCTATCTTCC 254
Db 30931 ATAAACGAATCTCTGACCTTCTCGCCCTATTTCACTGCTCTTATCTCTCC 30989

RESULT 4
AL773551/c 76225 bp DNA linear PRI 09-AUG-2002
LOCUS Human DNA sequence from clone Xbac-230F21 on chromosome 6,
DEFINITION complete sequence.
ACCESSION AL773551
VERSION AL773551.4 GI:22204642
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 76225)
Almeida, J.
DIRECT SUBMISSION
Submitted (01-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Aug 11, 2002 this sequence version replaced gi:21621738.

COMMENT

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M3 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Emi, EMBL; Sw,
SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs constructed by
the MHC HaploTYPE Consortium and collaborators. Further information
can be found at
<http://www.sanger.ac.uk/HGP/Chr6/MHC>
XXbac-230F21 is from a DNA-arts QBL human bac library VECTOR:
pBelobac11.

FEATURES

source Location/Qualifiers
1..76225
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="XXbac-230F21"
/clone_1id="DNA-arts-BAC.1-QBL.1"

ORIGIN

Query Match 29.7%; Score 89.4; DB 9; Length 76225;
Best Local Similarity 73.7%; Pred. No. 6.7e-14;
Matches 132; Conservative 0; Mismatches 36; Indels 11; Gaps 1;

Qy 87 TGTGATATATGATATATATAGCTTTTGTCCATGTTCTCGGCTTATAAACCCTC 146
Db 59989 TGTGCTATATATATATATAGCTTTTGTCCATGTTCTCGGCTTATAAACCCTC 59939
Qy 147 ACCCTGTTACAGCATTTGTTATAGGTTGAGTGGCTTCAAGACAAACTC- 205
Db 59988 GCCCTGTTACCATCTGTTGTTATAGTGGCTTCAAGACCAAGAGCTCA 59879
Qy 206 -----TCTCTCAACCTTCTCCAGCCCTCTGCTCTGCAACCTATCTTCC 254
Db 59878 ATAAACGAATCTCTGACCTTCTCGCCCTATTTCACTGCTCTTATCTCTCC 59820

RESULT 5
AB023050 112018 bp DNA linear PRI 22-NOV-2000
LOCUS Homo sapiens genomic DNA, chromosome 6p21.3, HLA class I region,
DEFINITION clone:832F2, complete sequence.
ACCESSION AB023050
VERSION AB023050.1 GI:5672605
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 112018)
Shihina, T., Tamiya, G., Oka, A., Takishima, N., Yamagata, T.,
Kikkawa, E., Iwata, K., Tomizawa, M., Okui, N., Kuwano, Y.,
Watanabe, K., Fukuzumi, Y., Itakura, S., Sugawara, C., Ono, A.,
Yamazaki, M., Taahiro, H., Ando, A., Iemura, T., Soeda, E., Kimura, M.,
Bahram, S. and Inoko, H.
Molecular dynamics of MHC genesis unraveled by sequence analysis of
the 1,796,938-bp HLA class I region
Proc. Natl. Acad. Sci. U.S.A. 96 (23), 13282-13287 (1999)
JOURNAL MEDLINE
PUBMED 20027539
10557312
2 (bases 1 to 112018)
Keicho, N., Ohashi, J., Tamiya, G., Nakata, K., Taguchi, Y., Azuma, A.,

TITLE
Ohishi,N., Emi,M., Park,M.H., Imoko,H., Tokunaga,K. and Kudoh,S.
Fine localization of a major disease-susceptibility locus for
diffuse panbronchiolitis
Am. J. Hum. Genet. 66 (2), 501-507 (2000)

JOURNAL
MEDLINE
20143375
10677310
3 (bases 1 to 112018)

REFERENCE
Shima,T. and Takishima,N.
Direct Submission
Submitted (29-JUN-1999) Takashi Shima, Tokai University School of
Medicine, Department of Molecular Life Science 2; Bohseidai,
Isehara, Kanagawa 259-1193, Japan
(E-mail: tshima@is.icc.u-tokai.ac.jp, Tel:81-463-93-1121,
Fax:81-463-94-8884)

FEATURES
source
Location/Qualifiers
1. 112018
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
/clone="833F2"

ORIGIN
Query Match 29.7%; Score 89.4; DB 9; Length 112018;
Best Local Similarity 73.7%; Pred. No. 6,5e-14;
Matches 132; Conservative 0; Mismatches 36; Indels 11; Gaps 1;

QY 87 TGTGATATTGATATATATAGAGTTTGTCCATAGTCTCGCTTAAACCCCTC 146
DB 111659 TGTGCTATTATATATATATAGTTTATGTCATGCTTCGCTTAAATTCCTCATT 111718
QY 147 ACCCTTGTACAGCATTTGTTATAGTTGATGTTTGGCTTGAAGAGCAAACTC- 205
DB 111719 GCCCTGTACCATCTGTTGTTAATGTTGGGCTTTAAGCCTCAGAGAGGTCTCA 111778
QY 206 -----TCTCTTCACCTTCTCCAGCCCTCTCTCTCTGACACCTCATCTTCCC 254
DB 111779 ATAAACGGAATCTCTGACCTTCTCTGACCTCTCTCTCTCTCTCTCTCTCTCC 111837

RESULT 6
AL662797/c 171627 bp DNA linear PRI 25-APR-2002
LOCUS Human DNA sequence from clone Xbac-252P9 on chromosome 6, complete
sequence.
ACCESSION AL662797
VERSION AL662797.7 GI:20338494
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 171627)
Almeida,J.
Direct Submission
Submitted (25-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 29, 2002 this sequence version replaced gi:20338227.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following

COMMENT

abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Emi, EMBL; Swi,
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs constructed by
the MHC Haplotype Consortium and collaborators. Further information
can be found at
http://www.sanger.ac.uk/HGP/Chre/MHC
Xbac-252P9 is from a CHORI-501 human bac - PGF cell line library
VECTOR: PHARAC2.1

FEATURES
source
Location/Qualifiers
1. 171627
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="Xbac-252P9"
/clone_id="CHORI-501"

ORIGIN
Query Match 29.7%; Score 89.4; DB 9; Length 171627;
Best Local Similarity 73.7%; Pred. No. 6,4e-14;
Matches 132; Conservative 0; Mismatches 36; Indels 11; Gaps 1;

QY 87 TGTGATATTGATATATATAGAGTTTGTCCATAGTCTCGCTTAAACCCCTC 146
DB 128619 TGTGCTATTATATATATATAGTTTATGTCATGCTTCGCTTAAATTCCTCATT 128560
QY 147 ACCCTTGTACAGCATTTGTTATAGTTGATGTTTGGCTTGAAGAGCAAACTC- 205
DB 128559 GCCCTGTACCATCTGTTGTTAATGTTGGGCTTTAAGCCTCAGAGAGGTCTCA 128500
QY 206 -----TCTCTTCACCTTCTCCAGCCCTCTCTCTCTGACACCTCATCTTCCC 254
DB 128499 ATAAACGGAATCTCTGACCTTCTCTGACCTCTCTCTCTCTCTCTCTCTCTCC 128441

RESULT 7
AL662848/c 185617 bp DNA linear PRI 24-APR-2002
LOCUS Human DNA sequence from clone Xbac-111D4 on chromosome 6, complete
sequence.
ACCESSION AL662848
VERSION AL662848.6 GI:19031747
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 185617)
Tracey,A.
Direct Submission
Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 28, 2002 this sequence version replaced gi:1947645.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Emi, EMBL; Swi,
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; information on the WORMPEP

COMMENT

/gene="PLOT1"
/number=10
111624, 111761
/gene="PLOT1"
/number=11
112003, 112167
/gene="PLOT1"
/number=12
114044, 114460
/gene="PLOT1"
/number=13
complement (114197, 114328)
/standard.name="D6s2115"
/note="SHGC-13451; The location is between each flanking
site of PCR primers."
/db_xref="GDB:738998"
complement (114300, 114430)
/note="WI-18721; The location is between each flanking site
of PCR primers."
/db_xref="GDB:4574312"
117236, 117362
/note="SGC3496; The location is between each flanking site
of PCR primers."
/db_xref="GDB:4580073"
complement (117712, 117836)
/standard.name="D6s2041"
/note="SHGC-10364; The location is between each flanking
site of PCR primers."
/db_xref="GDB:735115"
join (117792, 118849, 119161, 119271, 119544, 119652,
121626, 121682)
/gene="TUBB"
complement (join (117792, 118849, 119161, 119271,
119544, 119652, 121626, 121682))
/gene="TUBB"
complement (join (117792, 118849, 119161, 119271,
119544, 119652, 121626, 121682))
/gene="TUBB"
/product="Beta-tubulin"
/codon.start=1
/protein.id="BA863321.1"
/db_xref="GI:15277228"
/translation="MREIVHIOAGCCNQIGAFWEVISDEHIGIDPTGTHGSDIQL
KRISVYNEATGKCYPRALIVLEPGTMDSVSGPGQIFRDNVFVQSGAGNWA
DRIATGEGALVSDIVARKASCDLQGFOLHSIGSGTGSGMGILLISKIREY
DRIMTFEVSVPKSDTVVEPVATLSVQLVNDETCINEMALYDICEFTLKT
TPYGDINHVSATMSGVTTCTPRPOLNMLKLIANNVPPRLFPMPGAPLTSR
GSQQRALIVPELTQVFDKNNMAACDPHGRILYTAAYFRMRMKEVDEGMATVQ
NKSSYFVEMIPNNKTAACDIPRGKMAVTIGNSTALQDFKSISEFTMFRK
AFILMYEGMDMEFTEASNNMDLVSEYQYQDAVAEEEDFGSEAEBA"
117792, 118849
/gene="TUBB"
/number=4
complement (117961, 118093)
/gene="TUBB"
/note="SGC32809; The location is between each flanking site
of PCR primers."
/db_xref="GDB:4584619"
119161, 119271
/gene="TUBB"
/number=3
119544, 119652
/gene="TUBB"
/number=2
121626, 121682
/gene="TUBB"
/number=1
127010, 142364
/gene="K1A0170"
127010, 127148
/gene="K1A0170"
/number=1
join (127013, 127148, 128005, 128385, 128471, 128540,

128834, 130314, 130462, 130521, 130684, 130776,
133831, 134625, 134734, 134804, 136089, 138566,
138650, 138769, 138901, 139083, 139310, 139435,
139524, 139634, 141540, 141707)
/gene="K1A0170"
/note="Homologue to Drosophila photoreceptor protein
calphodien"
/codon.start=1
/protein.id="BA863322.1"
/db_xref="GI:15277229"
/translation="MEPTQAIWDVDEBEETSSSRGNVGRHIFSGAHPE
KDFPLHGXNVGMPCSVLPFSSIKQHAIELAKAPILNDGSLNKTQIR
PKYLSGVSHRLDDELIFADLLQYHLDVLPFSSGPTVBEPTVQGETPOQ
RLIAEDGSEEDVLSERRMVKRSSTSSSVIPEDEBESFVLGGPFPFAFNLS
DTVEEGQAPTEBASAAARGATVEAKGSAVYVEIQLEKQPIYKERNDPTKYR
GAGNGVPACVTLERSPPEDSDPTVDSDRPPRPAVYHRAQDPFGIDSDAE
EERIPATVPIPKRKIFHGVGTGGAGALHLOESQSGSDTVEBGRAPAVPLE
KQASWVINDTDEEVSALTLALHKSQPAIMWRDAEDMPQVRLQNSQTTTE

Query Match 29.7%; Score 89.4; DB 9; Length 200000;
Best Local Similarity 73.7%; Pred. No. 6.4e-14;
Matches 132; Conservative 0; Mismatches 36; Indels 11; Gaps 1;

QY 87 TGTGATATTATGATATATATATAGTTTGTCCATAGTTCTGGCTTATAACCCCTC 146
Db 36630 TGTGCTATTATATTATATATAGTTTGTCCATAGTTCTGGCTTATAATCCATT 36689

QY 147 ACCCTTGTAAGCATTTGTTATTAAGCTGATGTTTGGCTCGAAGCAAAATC- 205
Db 36690 GCCCTTGTAAGCATCTGTTGTTAATGTTGGGCTTTAAGCTCGAAGCAAGTCTCA 36749

QY 206 -----TCTCTCAACCTCTCCAGCCCTCCTGCTGACCTCATTTCTTCC 254
Db 36750 ATTAACGAATCTCTGACCTCTCCGCTCATTTCACTGCTCTTATCTTCC 36808

RESULT 9
AB100084 300050 bp DNA linear PRI 28-JUN-2003
LOCUS Pan troglodytes DNA, major histocompatibility complex class I
DEFINITION Pan troglodytes DNA, major histocompatibility complex class I
region, section 3/6.
ACCESSION AB100084 BA000041
VERSION AB100084.1 GI:32127780

KEYWORDS Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.

REFERENCE 1 Anzai, T., Shitina, T., Kimura, N., Yanggiya, K., Kohara, S.,
Shigenari, A., Yamagata, T., Kulski, J. K., Naruse, T. K., Fujimori, Y.,
Fukuzumi, Y., Yamazaki, M., Tashiro, H., Iwamoto, C., Umehara, Y.,
Imanishi, T., Meyer, A., Ikeo, K., Gojobori, T., Bahram, S. and Inoko, H.

TITLE Comparative sequencing of human and chimpanzee MHC class I regions
unveils insertions/deletions as the major path to genomic
divergence
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7708-7713 (2003)
MEDLINE 22709134
PUBMED 12799463

REFERENCE 2 (bases 1 to 300050)
AUTHORS Anzai, T. and Shitina, T.
TITLE Direct Submission
SUBMITTED (15-JAN-2003) Tatsuya Anzai, Tokai University School of
Medicine, Department of Molecular Life Science 2; Bohseidai,
Isehara, Kanagawa 259-1193, Japan
(E-mail: anzai@is.tcu.u-tokai.ac.jp, Tel: 81-463-93-1121 (ex. 2652),
Fax: 81-463-94-8884)

FEATURES
source 1..300050
Location/Qualifiers
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
misc_feature 1..6411

```

Gene .gene
    /note="Library: RPCI-43, Clone: 146D08"
    complement(1833..7664)
    /gene="rPIIH"
    complement(1833..2134)
    /gene="rPIIH"
    /number=14
    complement(join(1962..2134,2558..2636,2788..2835,
3488..3618,3800..3932,4179..4262,4447..4815,4677..4788,
5005..5093,5185..5281,5883..6014,6407..6511,6773..6909)))
    /gene="rPIIH"
    /product="transcription factor II H"
    /protein_id="BAC78171.1"
    /db_xref="gi:32127781"
    /translation="MESTPSRLNRYHLCQCNLOEFLGLSPGVLDRLYGHPTCLAV
FEEPLIANQWYRMFLTEQPLPOAVALTWKESKQAESETGLSGRLWHQOLP
GGIQLINPIFRONRILALGGKAMSDTSLQLPDGHARDVPSLDYAEERNEVVL
HNMVSPKSAVSODLALQLTSGAGKMSSTEPGSPCITSGAQPFLLDTPALWFMLO
YIQATGSRBMVDYEILTSFGSLTGDYVEGMSDSLNFLOHLERFGLVQRER
KRRRYITPLAINLSGVSAGAGTVHQPEFIVETNRYLVATSELDIALIAFSEM
LVRFEPMVAQVQRESVOQALASGITAQOIHFLRTRAHPWLKQTPVLPPTIIDQIR
LWELEEDRFTFEGVLYNQFLSQVDPELLAHARLGYLVFENSARKRLMVTTPGHSD
VRFVFRGHSS"
    complement(2558..2636)
    /gene="rPIIH"
    /number=13
    complement(2788..2835)
    /gene="rPIIH"
    /number=12
    complement(3488..3618)
    /gene="rPIIH"
    /number=11
    complement(3800..3932)
    /gene="rPIIH"
    /number=10
    complement(4179..4262)
    /gene="rPIIH"
    /number=9
    complement(4447..4515)
    /gene="rPIIH"
    /number=8
    complement(4677..4788)
    /gene="rPIIH"
    /number=7
    complement(5005..5093)
    /gene="rPIIH"
    /number=6
    complement(5185..5281)
    /gene="rPIIH"
    /number=5
    complement(5883..6014)
    /gene="rPIIH"
    /number=4
    complement(6407..6511)
    /gene="rPIIH"
    /number=3
    complement(6773..6912)
    /note="Library: RPCI-43, Clone: 158N24"
    complement(6773..6912)
    /gene="rPIIH"
    /number=2
    complement(7543..7664)
    /gene="rPIIH"
    /number=1
    complement(16929..27490)
    /gene="DDR"
    complement(join(16929..17069,17188..17337,17918..18152,
18620..18838,19076..19203,19352..19596,20702..20812,
21546..21711,22790..22941,23050..23145,23671..23917,
24026..24212,24743..24842,25102..25249,26790..27006,
27210..27312,27406..27490))
    /gene="DDR"
    /codon_start=1

```

| | |
|------|---|
| | /product="receptor tyrosine kinase" /protein_id="BAC78172.1" /db_xref="GI:32127782" |
| exon | /translacion="MGPEALSSLLILLLVASGDAMKGFHPDAKRCALGMQDRTPD SDISASSMSNSTAARHSDDGAMCPAGSVFPXEEELYQVLDLQRLLVALVMVGOGH AGGLDKSRSRYLRYSRDGRMMGMGDMRGOEVIISGNDEEGVLLKDLPMPYARLV PVPAPADRVSMVCRLRELYGCLMRDGILSYTA PVGOMTYSEAVYLNDSDYDGTAVGG LOYGGLGOLAGVGVGLDPFRKSQELRWPGDYIVGNSHSBSGCVYEMEEFPRIRAF QAMOVHCNNMHTLARLPFGVECFRRRGPAWANGEEPWRHNLSNLDPPARAIVPL GARVVRLOCFEFLFAGPWLLFSFISFDVVNNSPALGGTFPPAPMWPGPPPTNFS SLEETPRPSVDGDTILLNRSGRPBPPOEPBRPNPSPASCVNGSALLSNPA LEEETHVASVAGDTILLNRSGRPBPPOEPBRPNPSPASCVNGSALLSNPA RYLLIATYARP RPGRGPPTPMACPNTQAISGDYMEPEKGAILP PPNONSYPHYA EADIITVLGVGSNTAYVAPALP PGAVGDGPPKRVDFPKSRLLFEKTUSGGQGEVHLCE VDSDDIDLVDPFNVRKGPHLLVAVKILREPDAFKARNDFLEVKIMSLKUPNIIR LLVGCVQDDPLCMITDYMENGDLNQFLSAHQLEBKAEGA PGDOAAQGCTTISPMILL HVAADIGSVGYLATINVFHRDLTRNCIVGENFTIKIADFGMSRNLYAGDYRVOGR AVLP RRMWAGCITLMGKFTTSADVMAFGVTILEMYLMLCRAPFGQLTDDEVENAGEF FRDQRYVYLSRFPACPEGGLYEIMLRCKSRSEBORPFESQLHRLIADALNTV" complement (<16929 . 17069) |
| exon | /gene="DDR" /number=17 complement (17188 . 17337) /gene="DDR" |
| exon | /number=16 complement (17918 . 18152) /gene="DDR" /number=15 complement (18650 . 18836) /gene="DDR" |
| exon | /number=14 complement (19076 . 19203) /gene="DDR" /number=13 complement (19352 . 19596) /gene="DDR" |
| exon | /number=12 complement (20702 . 20812) /gene="DDR" /number=11 complement (21546 . 21711) /gene="DDR" |
| exon | /number=10 complement (22790 . 22941) /gene="DDR" /number=9 complement (23050 . 23145) /gene="DDR" |
| exon | /number=8 complement (23671 . 23917) /gene="DDR" /number=7 complement (24026 . 24212) /gene="DDR" |
| exon | /number=6 complement (24743 . 24842) /gene="DDR" /number=5 complement (25102 . 25249) /gene="DDR" |
| exon | /number=4 complement (26790 . 27006) /gene="DDR" /number=3 complement (27210 . 27312) /gene="DDR" |
| exon | /number=2 complement (27406 . 27490) /gene="DDR" /number=1 171051 . 172310 /gene="TEX-1" |
| CDS | join(171051 . 171260,171373 . 171633) |

/gene="IEX-1"
 /note="PRG1
 expressed in pancreatic carcinoma"
 /codon_start=1
 /protein_id="BAC78173.1"
 /db_xref="GI:32127783"
 /translation="MCHSRSCHPMTLILQAPAPSTNPGRRSGSEIFETDPLPEP
 AAPAGRPASRGHRRSRVLYPCVVRRLPVEEPNPAKRLFLTLTIVFOILMAE
 EGVPAFLPPEDANASLAPTPVSPVLEPNNLSESDVALDLSTLQHPAAF"
 <171051..171260
 /gene="IEX-1"
 /note="PRG1"
 /number=1
 171373..172310
 /gene="IEX-1"
 /note="PRG1"
 /number=2
 172876..187775
 /gene="HLOT1"
 gene

Query Match 29.2%; Score 87.8; DB 9; Length 300050;
 Best Local Similarity 73.2%; Pred. No. 1,7e-13;
 Matches 131; Conservative 0; Mismatches 37; Indels 11; Gaps 1;

QY 87 TGTGATATTGATATATATATAGTTTGTTCATAGTCCGCTTATTAACCCCTC 146
 Db 110384 TGTGCTATTATATATATATATAGTTTGTTCATAGTCCGCTTATTAATTCATT 110443
 QY 147 ACCCTGTTACAGCATTTGTTATAGTGTGATGTTAGGCTCAGAACAACTC- 205
 Db 110444 GCCCTGTTACCGCTGTGTGTTGATAGTGGGCTTAAAGCTCAGAACAGGTCTCA 110503
 QY 206 -----TCTCTCTACCTTCTCCAGCCCTCTCTCTCTGACACCTCATTTCTCC 254
 Db 110504 ATTAACGCAATCTCTCATTTCTCTCCGCTCATTTCAATCTCTCTTATCTCTCC 110562

RESULT 10
 HSPA20F2/c 481 bp DNA linear PRI 23-AUG-1996
 LOCUS H.sapiens flow-sorted chromosome 6 HindIII fragment, SC6PA20F2.
 DEFINITION Z79016
 VERSION Z79016.1 GI:1508294
 KEYWORDS Anonymous marker; single read.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 481)
 Mungall, A.J., Huckle, E., Langford, C., Ross, M.T. and Rice, C.M.
 TITLE Direct Submission
 JOURNAL Submitted (22-AUG-1996) The Sanger Centre, Wellcome Trust Genome
 Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
 humquerry@sanger.ac.uk
 Vector: pBS1ISK+.

COMMENT
 FEATURES
 source
 Location/Qualifiers
 1..481
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="SC6PA20F2"
 /sex="female"
 /tissue_type="EBV lymphoblastoid cell line"
 /clone_lib="SC6PA"
 /dev_stage="adult"
 /note="The estimated purity of the flow-sorted chromosome
 6 library is >97%"

ORIGIN
 Query Match 27.8%; Score 83.8; DB 9; Length 481;
 Best Local Similarity 76.3%; Pred. No. 2.9e-12;
 Matches 103; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 87 TGTGATATTGATATATATATAGTTTGTTCATAGTCCGCTTATTAACCCCTC 146
 Db 234 TGTGCTATTATATATATATATAGTTTGTTCATAGTCCGCTTATTAATTCATT 175
 QY 147 ACCCTGTTACAGCATTTGTTATAGTGTGATGTTAGGCTCAGAACAACTC 206
 Db 174 GCCCTGTTACCGCTGTGTGTTGATAGTGGGCTTAAAGCTCAGAACAGGCTCT 115
 QY 207 CTCTCTACCTTCTC 221
 Db 114 GACCTTCTCTGCCC 100

RESULT 11
 HSDJ93K22 113804 bp DNA linear PRI 04-MAR-2003
 LOCUS Human DNA sequence from clone RPI-93K22 on chromosome 6q14.1-15.3,
 complete sequence.
 DEFINITION AL050333
 ACCESSION AL050333.18 GI:6911641
 VERSION HTG.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 113804)
 Tracey, A.
 Direct Submission
 Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquerry@sanger.ac.uk
 On Feb 7, 2000 this sequence version replaced gi:6807622.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
 on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 This sequence was generated from part of bacterial clone contigs of human
 chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr6
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquerry@sanger.ac.uk

COMMENT
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest, except on the rare
 occasion of the clone being a YAC.
 RPI-93K22 is from the library RPI-1 constructed by the group of
 Piter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pCYPAC2.

FEATURES
 source
 Location/Qualifiers
 1..113804
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="6"
 /map="q14.1-15.3"

ORIGIN /clone="RP1-93K22"
/clone_id="RPCT-1"

Query Match 27.8%; Score 83.8; DB 9; Length 113804;
Best Local Similarity 76.3%; Pred. No. 2.2e-12;
Matches 103; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 87 TGTGATATATATATATATATATAGCTTTTGTCCATAGTTCCTGCTTAATAAACCCCTC 146
DB 112699 TGTGATATATATATATATATATAGCTTTTGTCCATAGTTCCTGCTTAATAAACCCCTC 112758

QY 147 ACCCTGTATACAGCATTTGTATATAGGTGATGCTTGGCTCAGAACCAAACTCT 206
DB 112759 GCCCTGTATACAGCATTTGTATATAGGTGATGCTTGGCTCAGAACCAAACTCT 112818

QY 207 CTCTCTACCTCTC 221
DB 112819 GACCTTCTCTGCCC 112833

RESULT 12

AC025610/c 170877 bp DNA linear HTG 30-MAR-2000
LOCUS Homo sapiens chromosome 6 clone RP11-413M1 map 6, WORKING DRAFT
DEFINITION SEQUENCE, 36 unordered pieces.

AC025610
AC025610.2 GI:7342184
HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
1 (bases 1 to 170877)
Homo sapiens chromosome 6, clone RP11-413M1
Unpublished
2 (bases 1 to 170877)
Britten, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bada, F.,
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choquet, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,
Dodgson, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Gallagan, J., Gardyna, S., Glend, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karatas, A.,
Klein, J., Labocque, K., Lamazares, R., Landers, T., Lehotzky, V.,
Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McPherson, L.,
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenka, J., Morrow, J.,
Murphy, T., Naylor, T., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testa, S., Theodore, J., Tirrell, A., Travers, M., Triggiano, J.,
Vasilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT
Submitted (12-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 30, 2000 this sequence version replaced gi:7229942.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/JM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
Project Information
Center project name: L8333

Center clone name: 413.M.1
----- Summary Statistics
Sequencing vector: M13; 877815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 147503 bases at least Q40
Consensus quality: 159832 bases at least Q30
Consensus quality: 164267 bases at least Q20
Insert size: 186000; agarose-fp
Insert size: 167377; sum-of-ctnigs
Quality coverage: 2.8 in Q20 bases; agarose-fp
Quality coverage: 3.1 in Q20 bases; sum-of-ctnigs

NOTE: This is a 'working draft' sequence. It currently
* consists of 36 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

| | | | |
|---|-------|--------------------------|-----------|
| 1 | 1682 | contig of 1682 bp | in length |
| * | 1683 | gap of 100 bp | |
| * | 1783 | 2939: contig of 1157 bp | in length |
| * | 2940 | 3039: gap of 100 bp | |
| * | 3040 | 4951: contig of 1312 bp | in length |
| * | 4952 | 5051: gap of 100 bp | |
| * | 5052 | 6630: contig of 1579 bp | in length |
| * | 6631 | 6730: gap of 100 bp | |
| * | 6731 | 8341: contig of 1611 bp | in length |
| * | 8342 | 8441: gap of 100 bp | |
| * | 8442 | 10096: contig of 1655 bp | in length |
| * | 10097 | 10196: gap of 100 bp | |
| * | 10197 | 11547: contig of 1351 bp | in length |
| * | 11548 | 11647: gap of 100 bp | |
| * | 11648 | 13051: contig of 1404 bp | in length |
| * | 13052 | 13151: gap of 100 bp | |
| * | 13152 | 14432: contig of 1281 bp | in length |
| * | 14433 | 14532: gap of 100 bp | |
| * | 14533 | 16465: contig of 1933 bp | in length |
| * | 16466 | 16565: gap of 100 bp | |
| * | 16566 | 18763: contig of 2198 bp | in length |
| * | 18764 | 18863: gap of 100 bp | |
| * | 18864 | 20876: contig of 2013 bp | in length |
| * | 20877 | 20976: gap of 100 bp | |
| * | 20977 | 23264: contig of 2288 bp | in length |
| * | 23265 | 23364: gap of 100 bp | |
| * | 23365 | 25715: contig of 2351 bp | in length |
| * | 25716 | 25815: gap of 100 bp | |
| * | 25816 | 29287: contig of 3472 bp | in length |
| * | 29288 | 29387: gap of 100 bp | |
| * | 29388 | 32438: contig of 3051 bp | in length |
| * | 32439 | 32538: gap of 100 bp | |
| * | 32539 | 33835: contig of 3257 bp | in length |
| * | 33836 | 35935: gap of 100 bp | |
| * | 35936 | 39035: contig of 3100 bp | in length |
| * | 39036 | 39135: gap of 100 bp | |
| * | 39136 | 40320: contig of 1185 bp | in length |
| * | 40321 | 40420: gap of 100 bp | |
| * | 40421 | 45654: contig of 5234 bp | in length |
| * | 45655 | 45754: gap of 100 bp | |
| * | 45755 | 50391: contig of 4637 bp | in length |
| * | 50392 | 50491: gap of 100 bp | |
| * | 50492 | 54449: contig of 3958 bp | in length |
| * | 54449 | 54549: gap of 100 bp | |
| * | 54550 | 59425: contig of 4876 bp | in length |
| * | 59426 | 59525: gap of 100 bp | |
| * | 59526 | 65171: contig of 5646 bp | in length |
| * | 65172 | 65271: gap of 100 bp | |
| * | 65272 | 68021: contig of 3750 bp | in length |
| * | 68022 | 69121: gap of 100 bp | |
| * | 69122 | 75111: contig of 5990 bp | in length |
| * | 75112 | 75211: gap of 100 bp | |

```
misc_feature 1., .1682
misc_feature /note="assembly_fragment" 1783., .2939
misc_feature /note="assembly_fragment" 3040., .4951
misc_feature /note="assembly_fragment" 5052., .6630
misc_feature /note="assembly_fragment" 6731., .8341
misc_feature /note="assembly_fragment" 8442., .10096
misc_feature /note="assembly_fragment" 10197., .11547
misc_feature /note="assembly_fragment" 11648., .13051
misc_feature /note="assembly_fragment" 13152., .14432
misc_feature /note="assembly_fragment" 14533., .16465
misc_feature /note="assembly_fragment" 16566., .18763
misc_feature /note="assembly_fragment" 18864., .20876
misc_feature /note="assembly_fragment" 20977., .23264
misc_feature /note="assembly_fragment" 23365., .25715
misc_feature /note="assembly_fragment" 25816., .29287
misc_feature /note="assembly_fragment" 29388., .32438
misc_feature /note="assembly_fragment" 33539., .35835
misc_feature /note="assembly_fragment" 35936., .39035
misc_feature /note="assembly_fragment" 39136., .40320
misc_feature /note="assembly_fragment" 40421., .45654
misc_feature /note="assembly_fragment" 45755., .50391
misc_feature /note="assembly_fragment" 50492., .54449
```

| | |
|------------|--|
| RESULT | 13 |
| AC021655 | |
| LOCUS | 143133 bp DNA linear HTG 21-SEP-2002 |
| DEFINITION | Homo sapiens chromosome 3 clone RP11-153K19, WORKING DRAFT SEQUENCE, 26 unordered pieces. |
| ACCESSION | AC021655 |
| VERSION | AC021655.21 GI:21431092 |
| KEYWORDS | HTG; HTGS PHASE1; HTGS_DRAFT. |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. |
| AUTHORS | Munzy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., I (bases 1 to 143133) |

Alsbrooks, S. L., Amaratunge, H. C., Aire, J. R., Ayer, M., Banks, ...
 Andarica, J., Benton, J., Bimace, K., Blankenburg, K., Bonnin, D.,
 Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N. P.,
 Bunay, C., Burch, P., Burkett, C., Burrell, K. L., Byrd, N. C.,
 Carron, T. F., Carter, M., Cavazos, S. R., Chacco, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,
 Cleveland, C. D., Cox, C., Coyle, M. D., Dathorne, S. R., David, R.,
 Davila, M. L., Davis, C., Davy-Carroll, L., Dedereich, D. A.,
 Delaney, K. R., Delgado, O., Denn, A. L., Ding, Y., Dinh, H. H.,
 Douthwaite, K. J., Draper, H., Dugan-Rocha, S., Durbin, K. J.,
 Earmhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foerster, P.,
 Gabis, J. A., Gao, Y., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
 Homisi, F., Howard, S., Huber, J., Holly, S., Hume, J., Jackson, L. E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvald, J., Kovar, C.,
 Kratoch, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, M., Louised, H.,
 Lozado, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maleshaari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mahoney, E., McLeod, M. P., Meador, M., Mei, G., Metzger, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, N., Morris, S.,
 Moser, M., Neal, D., Newton, S., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokwuko, S., Ogulu, M., Okunolu, G.,
 Oregunye, N., Oviedo, R., Pace, A., Paul, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Puy, L. L., Quiles, M., Ren, Y.,
 Rivers, M., Rojas, A., Rojoubokan, I., Rolfe, M., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shoostari, N., Sison, I.,
 Sodergren, E., Schnalke, T., Sparks, A., Stanley, H., Stone, H.,
 Suttor, A., Swatek, A., Tabor, P., Tamezisa, K., Tamezisa, K., Tang, H.,
 Tanney, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Umami, K., Vazquez, L., Vera, V., Villalon, D., Vinson, R., Wang, O.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wlaczek, R., Woodem, S., Wolley, K.,

FEATURES
source

```

/db_xref="caxon:9606"
/chromosome="3q"
/clone="RP11-776D7"

Query Match      27.4%; Score 82.6; DB 2; Length 163862;
Best Local Similarity 75.2%; Pred. No. 4,6e-12;
Matches 103; Conservative 0; Mismatches 34; Indels 0; Gaps 0

ORIGIN

Query 82 AGGGTTGATATATATGATATATATATAGTTTGTTCATAGTTCCTGGCTTATATACC 141
Db 4701 AGGATGTGTGTCTATGATATATATATGTTTATGTTCATAGTTCCTTCATATCTC 4766
Query 142 CCTCACCCTTGTATCAGTCATTTTGTATTAAGTTGATGTTTGGCCTCAGAACAA 201
Db 4761 CCATAGCCCTTGTATCAGTCATTTTGTATTAAGTTGATGTTTGGCCTCAGAACAA 4822
Query 202 ACTCTCTCTCCACCTT 218
Db 4821 ATCTCTCTGACCTCTCT 4837

RESULT 15
LOCUS AC026559 165082 bp DNA linear HTG 24-AUG-2002
DEFINITION Homo sapiens chromosome 4 clone RP11-776D7 map 4, WORKING DRAFT
SEQUENCE, 8 unordered pieces.
AC026559
AC026559.2 GI:7467245
HTG, HTGS_PHASE1, HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 165082)
Homo sapiens chromosome 4, clone RP11-776D7
Unpublished
2 (bases 1 to 165082)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,C., Baria,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campione,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Larocque,K., Lamazeres,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,G., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,H., McEwan,P., McGurk,A., McKernan,K., McPherson,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Menga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tassile,S., Theodore,J., Tittrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zaitoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 165082)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Baria,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campione,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Larocque,K., Lamazeres,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,G., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,H., McEwan,P., McGurk,A., McKernan,K., McPherson,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Menga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tassile,S., Theodore,J., Tittrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zaitoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 165082)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Baria,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campione,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Larocque,K., Lamazeres,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,G., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,H., McEwan,P., McGurk,A., McKernan,K., McPherson,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Menga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tassile,S., Theodore,J., Tittrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zaitoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 165082)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Baria,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campione,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Larocque,K., Lamazeres,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,G., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,H., McEwan,P., McGurk,A., McKernan,K., McPherson,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Menga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tassile,S., Theodore,J., Tittrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zaitoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 165082)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Baria,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campione,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Larocque,K., Lamazeres,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,G., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,H., McEwan,P., McGurk,A., McKernan,K., McPherson,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Menga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tassile,S., Theodore,J., Tittrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zaitoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 165082)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Baria,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campione,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Larocque,K., Lamazeres,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,G., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,H., McEwan,P., McGurk,A., McKernan,K., McPherson,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Menga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tassile,S., Theodore,J., Tittrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zaitoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 165082)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Baria,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campione,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Larocque,K., Lamazeres,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,G., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,H., McEwan,P., McGurk,A., McKernan,K., McPherson,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Menga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Strange
```


Klein, J., Laroque, K., Iamazaki, R., Ianders, T., Lehoczy, J.,
 Levine, R., Lien, G., Liu, G., Locke, K., MacDonald, P., Margis, N.,
 McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McPheters, R.,
 Meldrum, J., Meunier, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Olivieri, T.M., Oliver, J., Peterson, K., Pierre, N.,
 Pisanik, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
 Vassiliev, H., Viel, R., Vo, X., Wilson, B., Wu, X., Wyman, D., Ye, W., J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 330 Charles Street, Cambridge, MA 02141, USA
 On Apr 6, 2000 this sequence version replaced gi:1283372.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RV/RepeatMasker.html>

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 12:26:39 / Search time 120.53 seconds

(without alignments)
10609.014 Million cell updates/sec

Title: US-09-966-880a-10_COPY_6264_6564

Perfect score: 301
1 gctcaagaagacacttaga.....ctccacaagcgctcaag 301

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

```
1: Geneseq_29JAN04:*
2: geneseqn1980s:*
3: geneseqn1990s:*
4: geneseqn2000s:*
5: geneseqn2001as:*
6: geneseqn2001bs:*
7: geneseqn2002s:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2004s:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 301 | 100.0 | 6564 | 3 AAC55314 | AAC55314 Human act |
| 2 | 73.4 | 24.4 | 355 | 4 AAI89801 | AAI89801 Human pol |
| 3 | 71.8 | 23.9 | 612 | 5 ABV58443 | ABV58443 Human pro |
| 4 | 70.4 | 23.4 | 740 | 4 AAI96626 | AAI96626 Human neu |
| 5 | 68.8 | 22.9 | 864 | 4 AAI86523 | AAI86523 Human pol |
| 6 | 68.8 | 22.9 | 2102 | 4 AAH15837 | AAH15837 Human CDN |
| 7 | 67 | 22.3 | 875 | 4 AAK82024 | AAK82024 Human imm |
| 8 | 67 | 22.3 | 35465 | 4 AAF54723 | AAF54723 Nucleotid |
| 9 | 67 | 22.3 | 35465 | 7 ABT17380 | ABT17380 Human IG |
| 10 | 67 | 22.3 | 36991 | 7 ABT17382 | ABT17382 Human IG |
| 11 | 66.8 | 22.2 | 621 | 6 ABN62372 | ABN62372 Human can |
| 12 | 65.8 | 21.9 | 452 | 4 AAK62248 | AAK62248 Human imm |
| 13 | 65.8 | 21.9 | 1524 | 4 AAL02965 | AAL02965 Human rep |
| 14 | 65.8 | 21.9 | 1524 | 4 ABA07734 | ABA07734 Human cva |
| 15 | 65 | 21.6 | 452 | 8 ACH26427 | ACH26427 Human act |
| 16 | 65 | 21.6 | 2141 | 4 AAS02525 | AAS02525 Human sec |
| 17 | 65 | 21.6 | 2510 | 6 AAD31155 | AAD31155 Human Alz |
| 18 | 64.4 | 21.4 | 13069 | 4 AAK36350 | AAK36350 Human car |
| 19 | 64.4 | 21.4 | 13069 | 4 AAK79980 | AAK79980 Human imm |
| 20 | 64.4 | 21.4 | 13069 | 4 ADA77044 | ADA77044 Human car |
| 21 | 64.4 | 21.4 | 32203 | 4 AAS36349 | AAS36349 Human car |
| 22 | 64.4 | 21.4 | 32203 | 9 ADE47043 | ADE47043 Human car |
| 23 | 64.4 | 21.4 | 47513 | 4 AAK74979 | AAK74979 Human imm |

| | | | | | |
|----|------|------|--------|---------------|--------------------|
| 24 | 63.4 | 21.1 | 21313 | 4 AAK82710 | AAK82710 Human imm |
| 25 | 63.4 | 21.1 | 222930 | 6 ABK84349 | ABK84349 Human CDN |
| 26 | 62.6 | 20.8 | 110000 | 6 ABX08336_13 | Continuation (14 o |
| 27 | 62.6 | 20.8 | 110000 | 7 ACFA2745_0 | ACFA2745 Human ALM |
| 28 | 62.4 | 20.7 | 14800 | 7 ADC85848 | ADC85848 Human GPC |
| 29 | 62 | 20.6 | 957 | 3 AAC74307 | AAC74307 Human sec |
| 30 | 62 | 20.6 | 1198 | 4 AAI92788 | AAI92788 Human pol |
| 31 | 62 | 20.6 | 1733 | 3 AAC76894 | AAC76894 Human ORF |
| 32 | 61.8 | 20.5 | 736 | 4 AAK69223 | AAK69223 Human imm |
| 33 | 61.4 | 20.4 | 48727 | 4 AAK67375 | AAK67375 Human imm |
| 34 | 60.4 | 20.1 | 148567 | 7 ABSS5500 | ABSS5500 Gene enco |
| 35 | 60.4 | 20.1 | 148567 | 8 ACA62841 | ACA62841 Human kin |
| 36 | 59.2 | 19.7 | 967 | 6 ABZ11514 | ABZ11514 Human pol |
| 37 | 58.2 | 19.3 | 39776 | 7 AAL51353 | AAL51353 Human sec |
| 38 | 57.4 | 19.1 | 416 | 8 ACH48924 | ACH48924 Human leu |
| 39 | 57.4 | 19.1 | 585 | 4 AAI10627 | AAI10627 Human CDN |
| 40 | 57 | 18.9 | 2626 | 4 AAI17370 | AAI17370 Human CDN |
| 41 | 56.6 | 18.8 | 5879 | 6 ABL32268 | ABL32268 Human imm |
| 42 | 55.8 | 18.5 | 174493 | 7 ACA61395 | ACA61395 Novel hum |
| 43 | 55.8 | 18.5 | 174493 | 9 AAD59937 | AAD59937 Human kin |
| 44 | 55.4 | 18.4 | 381 | 6 ABV95367 | ABV95367 Human pan |
| 45 | 55.4 | 18.4 | 546 | 3 AAK31133 | AAK31133 Human col |

ALIGNMENTS

| | |
|---|--|
| RESULT 1 | |
| AAC55314 | |
| ID AAC55314 standard; DNA; 6564 BP. | |
| XX | |
| AC AAC55314; | |
| XX | |
| DT 05-FEB-2001 (first entry) | |
| XX | |
| DE Human activation-induced cytidine deaminase genomic DNA SEQ ID NO:10. | |
| XX | |
| KM Activation-induced cytidine deaminase; AID; cytidine deaminase; | |
| KM immune related disease; allergy; allergic disease; anti-allergic; | |
| KM antianemic; antiasthmatic; ophthalmological; anti-HIV; dermatological; | |
| KM gene therapy; B cell associated immune system disorder; food allergy; | |
| KM immunodeficiency disease; immunoglobulin A deficiency disease; asplenia; | |
| KM Iga nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis; | |
| KM drug allergy; allergic rhinitis; Rosen disease; Digorge disease; AIDS; | |
| KM ataxia telangiectasia; common variable immunodeficiency disorder; | |
| KM major histocompatibility class II deficiency disease; | |
| KM auto immunodeficiency syndrome; Igg subclases selection disorder; ds. | |
| XX | |
| OS Homo sapiens. | |
| XX | |
| PN WO200058480-A1. | |
| XX | |
| PD 05-OCT-2000. | |
| XX | |
| PF 28-MAR-2000; 2000WO-JP001918. | |
| XX | |
| PR 29-MAR-1999; 99JP-00087192. | |
| XX | |
| PR 24-JUN-1999; 99JP-00178999. | |
| XX | |
| PA 27-DEC-1999; 99JP-00371382. | |
| XX | |
| PA (NIBS) JAPAN TOBACCO INC. | |
| XX | |
| PI (HONJ/) HONJO T. | |
| XX | |
| PI Honjo T, Muramatsu M; | |
| XX | |
| DR WPI; 2000-611715/58. | |
| XX | |
| PT Nucleic acid encoding activation induced cytidine deaminase, useful as a | |
| PT target for drug development for immune-related diseases including | |
| PT allergies. | |
| XX | |
| PS Claim 17; Page 145-150; 174pp; Japanese. | |
| XX | |

| | | |
|-----------------------|---|----|
| PR | 13-DEC-2000; 2000US-0255281P. | XX |
| XX | | XX |
| PA | (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC. | XX |
| PI | Schlegel R, Endege WO, Monahan JE; | XX |
| DR | WPI; 2001-662795/76. | XX |
| XX | | XX |
| PT | Novel isolated nucleic acid molecule associated with cancerous state of | XX |
| PT | prostate cells and correlating with presence of prostate cancer, useful | XX |
| PT | for detecting presence of prostate cancer, stage of prostate cancer. | XX |
| PS | Claim 1; Page 11218; 11750pp; English. | XX |
| CC | | XX |
| CC | The invention relates to an isolated nucleic acid molecule (1) comprising | XX |
| CC | a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the | XX |
| CC | specification or its complement. (1) is useful for: (a) assessing whether | XX |
| CC | a patient is afflicted with prostate cancer; (b) monitoring the | XX |
| CC | progression of prostate cancer in a patient; (c) assessing the efficacy | XX |
| CC | of a test compound to inhibit prostate cancer in a patient; (d) assessing | XX |
| CC | the efficacy of a therapy for inhibiting prostate cancer in a patient; | XX |
| CC | (e) selecting a composition for inhibiting prostate cancer in a patient; | XX |
| CC | (f) assessing the prostate cell carcinogenic potential of a compound; (g) | XX |
| CC | determining whether prostate cancer has metastasized in a patient; (h) | XX |
| CC | assessing the aggressiveness or indolence of prostate cancer in a patient | XX |
| CC | ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker | XX |
| SQ | Sequence 612 BP; 157 A; 125 C; 124 G; 206 T; 0 U; 0 Other; | XX |
| | | XX |
| Query Match | 23.9%; Pred 71.8; DB 5; Length 612; | |
| Best Local Similarity | 69.8%; Score No. 1.3e-11; | |
| Matches | 97; Conservative 0; Mismatches 42; Indels 0; Gaps 0 | |
| Qy | 98 GATATATATATAGGTTTGTGCATAGTCTGCTTATTAACCCCGCTACCGCTTGTAC 157 | |
| Db | 204 GATACGATATAGGTTTGTGCCAGGTTCTGGCTCAACAATCCATAGCCCTGTAT 145 | |
| Qy | 158 AGTCATTGTATTAAGTTGGATGGTTTAGGCGCTCAAGAACAAACTCTCTCTCACCT 217 | |
| Db | 144 AGTTTTTGTATTAATGTTGGGTGTGTCGGGGCTCAGGAATATGATCTCTGACCTTC 85 | |
| Qy | 218 TCTCCAGCGCTCCTGCTC 236 | |
| Db | 84 TCTCCTCTTTTACCTGTC 66 | |
| | | |
| RESULT 4 | | |
| AA196626/c | | |
| AA196626 | standard; cDNA; 740 BP. | |
| AA196626; | | |
| DT | 13-NOV-2001 (first entry) | |
| DE | Human neuroblastoma expressed polynucleotide SEQ ID NO 2701. | |
| XX | | |
| XX | Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss. | |
| XX | | |
| XX | Homo sapiens. | |
| XX | | |
| XX | WO200166719-A1. | |
| PD | 13-SEP-2001. | |
| XX | | |
| PF | 02-MAR-2001; 2001MO-JP001629. | |
| XX | | |
| PR | 07-MAR-2000; 2000JP-00159195. | |
| XX | | |
| PA | (CHIB-) CHIBA PREFECTURE. | |
| PA | (HISM) HISAMITSU PHARM CO LTD. | |
| PI | Nakagawara A; | |
| XX | | |

DR WPI; 2001-565584/63.

XX Nucleic acids originating in gene expressed in human neuroblastoma,

PT useful as probe or primer in diagnosing prognosis of human neuroblastoma,

PT malignancy and susceptibility indicator or tumor marker for anti-cancer

PT agents.

XX

PS Claim 1; Page 1981; 2979pp; Japanese.

XX

CC The invention relates to novel genes (AA193926-AA197963) expressed in

CC human neuroblastoma. The nucleic acids are applicable as a probe or

CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and

CC susceptibility indicators or tumor markers for anti-cancer agents. The

CC gene information for diagnosing prognosis is related to factors similar

CC to that for N-myc and TrkA genes

XX

SQ Sequence 740 BP; 184 A; 122 C; 199 G; 218 T; 0 U; 17 Other;

QY Query Match 23.4%; Score 70.4; DB 4; Length 740;

Best Local Similarity 83.6%; Pred. No. 3,7e-11;

Matches 92; Conservative 0; Mismatches 16; Indels 2; Gaps 1

DB 188 GCCCTCAGAGCAAACTCTCTCTCAGCCCTTCAGCCCTCTGTCTGTGCACTCAT 247

562 GCCCCAGAGCAAAAGTTCTCTCTGTGACCTCTCCCTCCCTCTGTCTGTGCTTTTCAT 503

QY 248 TCTTCCCTGAGGC--CACATAGAAACTGAACTCTCTTTCACAAAGCGG 295

502 TCTCCCCCAAGGCTACCATAGAACTGAATCCCTCTTCCCAAGCAG 453

DB

RESULT 5

AA186523/c

ID AA186523 standard; cDNA; 864 BP.

XX

AC AA186523;

XX

DT 06-NOV-2001 (first entry)

XX

XX Human polynucleotide SEQ ID NO 6563.

XX

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

XX nervous system disorders; arthritis; inflammation; ss.

XX

OS Homo sapiens.

XX

PN WC020164835-A2.

XX

PD 07-SEP-2001.

XX

XX 26-FEB-2001; 2001WO-US004927.

XX

XX 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT;

XX

XX WPI; 2001-514838/56.

DR P-PSDB; AA006592.

XX

PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing

PT and treating e.g. leukemia, inflammation and immune disorders.

XX

PS Claim 1; SEQ ID NO 6563; 1399pp + Sequence Listing; English.

XX

CC The invention relates to human polynucleotides (AA179941-AA193841) and

CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

PR 14-AUG-2000; 2000US-0225214P

Matches 107; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 6 CTGGCAATATCATATACGTATAGGGTGTGTGATATTAAGAAATATATATAGATTTTGTG 118
Db 138 CCTCAATATCTTCCTTTTGGTTAGGCGCTGTGTATGATGACATATAGTTTGTGTC 187
QY 120 CATAGTTCCTGGCTTATTAACCCCTCCACCCCTGTGTACAGTCATTTGTTATTAAGTTGGA 179
Db 198 CATGTTCCCGGTTCAATACATCAATTCGCCCTGTGTGACGTCCTTTTGTATACCTGTGGG 257
QY 180 TGGTTTAGGCGCTCAGAAAGCAAACTCTCTCTCACTCTTCACAGCCCTCCTGT 223
Db 258 TGTGTTAGGCGCTCTCGGCAGGCGCTCTACACTTTCTTGCCCTTTCTTACTCT 311

RESULT 12
AAK62248
ID AAK62248 standard; cDNA; 452 BP

AC AAK62248;

DT 06-NOV-2001 (first entry)

| Accession | Gene Symbol | Gene Name |
|-----------|-------------|---|
| AF067892 | CDNA | Human immune/haematopoietic antigen encoding CDNA SEQ ID NO:7308. |

Human; immune; haematopoietic; immune/haematopoietic antigen; cancer
 KW
 cytostatic; gene therapy; vaccine; metastasis; ss.
 KW

OS Homo sapiens.

PN W0200157182-A2.

PD 09-AUG-2001
yy

PF 17-JAN-2001; 2001WO-US001354
XX

[illegible][illegible]

| | | | |
|----|-------------|------------------|--|
| PR | 08-SEP-2000 | 2000US-0232080P | |
| PR | 08-SEP-2000 | 2000US-0232081P | |
| PR | 12-SEP-2000 | 2000US-0231568P | |
| PR | 14-SEP-2000 | 2000US-0232397P | |
| PR | 14-SEP-2000 | 2000US-0232398P | |
| PR | 14-SEP-2000 | 2000US-0232399P | |
| PR | 14-SEP-2000 | 2000US-0232400P | |
| PR | 14-SEP-2000 | 2000US-0232401P | |
| PR | 14-SEP-2000 | 2000US-0233063P | |
| PR | 14-SEP-2000 | 2000US-0233064P | |
| PR | 14-SEP-2000 | 2000US-0233065P | |
| PR | 21-SEP-2000 | 2000US-0234423P | |
| PR | 21-SEP-2000 | 2000US-0234474P | |
| PR | 25-SEP-2000 | 2000US-0234997P | |
| PR | 25-SEP-2000 | 2000US-0234998P | |
| PR | 26-SEP-2000 | 2000US-0235584P | |
| PR | 27-SEP-2000 | 2000US-0235583AP | |
| PR | 27-SEP-2000 | 2000US-0235583BP | |
| PR | 29-SEP-2000 | 2000US-0236367P | |
| PR | 29-SEP-2000 | 2000US-0236368P | |
| PR | 29-SEP-2000 | 2000US-0236369P | |
| PR | 29-SEP-2000 | 2000US-0236370P | |
| PR | 02-OCT-2000 | 2000US-0237032P | |
| PR | 02-OCT-2000 | 2000US-0237033P | |
| PR | 02-OCT-2000 | 2000US-0237039P | |
| PR | 02-OCT-2000 | 2000US-0237040P | |
| PR | 13-OCT-2000 | 2000US-0239933P | |
| PR | 13-OCT-2000 | 2000US-0239937P | |
| PR | 20-OCT-2000 | 2000US-0240960P | |
| PR | 20-OCT-2000 | 2000US-0241212P | |
| PR | 20-OCT-2000 | 2000US-0241785P | |
| PR | 20-OCT-2000 | 2000US-0241786P | |
| PR | 20-OCT-2000 | 2000US-0241787P | |
| PR | 20-OCT-2000 | 2000US-0241809P | |
| PR | 20-OCT-2000 | 2000US-0241826P | |
| PR | 01-NOV-2000 | 2000US-0244617P | |
| PR | 08-NOV-2000 | 2000US-0246475P | |
| PR | 08-NOV-2000 | 2000US-0246476P | |
| PR | 08-NOV-2000 | 2000US-0246477P | |
| PR | 08-NOV-2000 | 2000US-0246478P | |
| PR | 08-NOV-2000 | 2000US-0246533P | |
| PR | 08-NOV-2000 | 2000US-0246534P | |
| PR | 08-NOV-2000 | 2000US-0246552P | |
| PR | 08-NOV-2000 | 2000US-0246556P | |
| PR | 08-NOV-2000 | 2000US-0246577P | |
| PR | 08-NOV-2000 | 2000US-0246582P | |
| PR | 08-NOV-2000 | 2000US-0246583P | |
| PR | 08-NOV-2000 | 2000US-0246609P | |
| PR | 08-NOV-2000 | 2000US-0246610P | |
| PR | 08-NOV-2000 | 2000US-0246611P | |
| PR | 17-NOV-2000 | 2000US-0249211P | |
| PR | 17-NOV-2000 | 2000US-0249212P | |
| PR | 17-NOV-2000 | 2000US-0249213P | |
| PR | 17-NOV-2000 | 2000US-0249214P | |
| PR | 17-NOV-2000 | 2000US-0249215P | |
| PR | 17-NOV-2000 | 2000US-0249216P | |
| PR | 17-NOV-2000 | 2000US-0249217P | |
| PR | 17-NOV-2000 | 2000US-0249218P | |
| PR | 17-NOV-2000 | 2000US-0249244P | |
| PR | 17-NOV-2000 | 2000US-0249245P | |
| PR | 17-NOV-2000 | 2000US-0249246P | |
| PR | 17-NOV-2000 | 2000US-0249265P | |
| PR | 17-NOV-2000 | 2000US-0249293P | |
| PR | 17-NOV-2000 | 2000US-0249295P | |
| PR | 17-NOV-2000 | 2000US-0249297P | |
| PR | 17-NOV-2000 | 2000US-0249299P | |

| | | |
|-----------------------|--|------------------|
| PR | 17-NOV-2000; | 2000US-0249300P. |
| PR | 01-DEC-2000; | 2000US-0250160P. |
| PR | 01-DEC-2000; | 2000US-0250391P. |
| PR | 05-DEC-2000; | 2000US-0251030P. |
| PR | 05-DEC-2000; | 2000US-0251988P. |
| PR | 05-DEC-2000; | 2000US-0256719P. |
| PR | 06-DEC-2000; | 2000US-0251479P. |
| PR | 08-DEC-2000; | 2000US-0251856P. |
| PR | 08-DEC-2000; | 2000US-0251868P. |
| PR | 08-DEC-2000; | 2000US-0251869P. |
| PR | 08-DEC-2000; | 2000US-0251989P. |
| PR | 08-DEC-2000; | 2000US-0251990P. |
| PR | 11-DEC-2000; | 2000US-0254097P. |
| PR | 05-JAN-2001; | 2001US-0259678P. |
| PA | (HUMA-) HUMAN GENOME SCI INC. | |
| XX | | |
| PI | Rosen CA, Barash SC, Ruben SM; | |
| XX | | |
| DR | WPI; 2001-465570/50. | |
| XX | | |
| PT | Isolated nucleic acid molecule encoding a reproductive system antigen is used in preventing, treating or ameliorating a medical condition. | |
| XX | | |
| PS | Disclosure; SEQ ID NO 5653; 1297bp + Sequence Listing; English. | |
| XX | | |
| CC | The present invention provides the protein and coding sequences of a number of human reproductive system related antigens. These can be used in the prevention and treatment of reproductive system disorders, including cancer. The present sequence is a genomic sequence encoding a protein of the invention | |
| CC | | |
| SQ | Sequence 1524 BP; 357 A; 369 C; 387 G; 411 T; 0 U; 0 Other; | |
| XX | | |
| Query Match | 21.9%; Score 65.8; DB 4; Length 1524; | |
| Best Local Similarity | 68.4%; Pred.No.1.3e-09; | |
| Matches | 91; Conservative 0; Mismatches 42; Indels 0; Gaps 0; | |
| QY | 89 TGATATTAGATATATATATATGTTTTCACATAGTCTCGTGCTTATAAACCCTCAG 148 1086 TGCTGTGTTTATGATAGACAGTTTCCCTCATGATCCCTGGCTCATTACTCCCATAGC 1145 | |
| DB | | |
| QY | 149 CCTGTGTACATCACTTTGCTTAAAGTTGATGGTTTAGGGCTAGAGAACAACCTCT 208 1146 CCTGTGTACACTCTTTGTTGTATATTGATGAGTGTTAGGCCACAGGAACGACTCT 1205 | |
| QY | 209 CTCACCTCTGC 221 | |
| DB | 1206 CCTCTCTCTGCC 1218 | |
| RESULT 14 | | |
| ABA07734 | | |
| ID | ABA07734 standard; DNA; 1524 BP. | |
| XX | | |
| AC | ABA07734; | |
| XX | | |
| DT | 11-JAN-2002 (first entry) | |
| XX | | |
| DE | Human ovarian and breast cancer associated polynucleotide SEQ ID NO 529. | |
| XX | | |
| KW | Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral; anticarcinogenic; hepatocytic; antidiabetic; antiinflammatory; antifungal; vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic; cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder; neurological disease; infection; human; secreted protein; ds. | |
| KM | | |
| OS | Homo sapiens. | |
| XX | | |
| XX | | |
| PN | WO200155325-A2. | |
| XX | | |
| FD | 02-AUG-2001. | |
| XX | | |

PT polypeptides, useful for diagnosing, treating and/or preventing human
PT diseases and disorders, particularly ovarian and/or breast cancer.
XX
PS Disclosure; SEQ ID NO 529; 577bp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABA07454-ABA08224) and proteins
CC (ABJ0743-ABJ0980) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer; e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC / (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIP0 at http://wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1524 BP; 357 A; 369 C; 387 G; 411 T; 0 U; 0 Other;
Query Match 21.9%; Score 65.8; DB 4; Length 1524;
Best Local Similarity 68.4%; Pred. No. 1.3e-09;
Matches 91; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 89 TGAATATGATATATATATAGTTTGTTCATAGTCTCGGCTATATAACCCCTCAC 148
Db 1086 TGCTGTGTGTATGATAGACATTTTCTCCATGATCTGCTCATTAATCCCATAC 1145
QY 149 CCTGTATACAGTATTTGTTATAGGTTGATGCTTACAGCCACAGAAACTCTCT 208
Db 1146 CCTGTATACAGTCTTTTGTGTTATATGATGCTGTTAGGCCACAGAAACAGACTCT 1205
QY 209 CTCTACCTTCTC 221
Db 1206 CCTCTCTCTGCC 1218
RESULT 15
ACH26427
XX ACH26427 standard; cDNA; 452 BP.
XX
AC ACH26427;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human adult ovary cDNA #4807.
XX
KM Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KM genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Dtmnac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX

DR WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 13639; 44p; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH2769-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 452 BP; 113 A; 109 C; 87 G; 129 T; 0 U; 14 Other;
Query Match 21.6%; Score 65; DB 8; Length 452;
Best Local Similarity 68.6%; Pred. No. 1.3e-09;
Matches 120; Conservative 0; Mismatches 50; Indels 5; Gaps 2;
QY 83 GGGTGTGATATATATATATATATAGTTTGTTCATAGTCTCGGCTATATAACCC 142
Db 125 GGGTGTGATATATATATATATATAGTTTGTTCATAGTCTCGGCTATATACTCC 182
QY 143 CCTACCCCTTGTATACATCTTTGTTATAGGTTGATGCTTACAGCCACAGAA---GCA 199
Db 183 CAAACCCCTTGTATAGCTTTTGTATATATATATATATAGGCTTACAGAAACAGC 242
QY 200 AAAGCTCTCTCTACCTCTTCCAGCCCTCTGCTCTGACCTCATTTCTTCC 254
Db 243 AGGCTGTTTCTCAGATCTTCTCCGACCTCTTTCACCTGCTGCTTTCTCC 297
Search completed: March 12, 2004, 18:07:59
Job time : 124.53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 17:39:25 ; Search time 22.3958 Seconds

(without alignments)
7458.540 Million cell updates/sec

Title: US-09-966-880A-10_COPY_6264_6564

Perfect score: 301 1 gcatcaagagacacttagta.....cttcacacagcggtcaag 301

Sequence: 1 gcatcaagagacacttagta.....cttcacacagcggtcaag 301

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgm2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgm2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgm2_6/prodata/2/ina/6A_COMB.seq.*
4: /cgm2_6/prodata/2/ina/6B_COMB.seq.*
5: /cgm2_6/prodata/2/ina/PCOTUS_COMB.seq.*
6: /cgm2_6/prodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Length | DB ID | Description |
|------------|-------|--------------|-----------------------|-------------------|
| 1 | 60.4 | 20.1 148567 | 4 US-09-801-876B-3 | Sequence 3, Appl |
| 2 | 60.4 | 20.1 148567 | 4 US-10-254-869-3 | Sequence 3, Appl |
| 3 | 55.8 | 18.5 174493 | 4 US-09-804-971A-3 | Sequence 3, Appl |
| 4 | 55.8 | 18.5 174493 | 4 US-10-238-709-3 | Sequence 3, Appl |
| 5 | 53.2 | 17.7 649 | 4 US-09-040-984-59 | Sequence 59, Appl |
| 6 | 53.2 | 17.7 649 | 4 US-09-123-912-59 | Sequence 59, Appl |
| 7 | 53.2 | 17.7 649 | 4 US-09-643-597-59 | Sequence 59, Appl |
| 8 | 53.2 | 17.7 649 | 4 US-09-480-884A-59 | Sequence 59, Appl |
| 9 | 53.2 | 17.7 649 | 4 US-09-542-615A-59 | Sequence 59, Appl |
| 10 | 53.2 | 17.7 649 | 4 US-09-606-421B-59 | Sequence 59, Appl |
| 11 | 53.2 | 17.7 649 | 4 US-09-221-107-59 | Sequence 59, Appl |
| 12 | 48.4 | 16.1 4080 | 4 US-09-016-434A-1342 | Sequence 1342, Ap |
| 13 | 48.4 | 16.1 246240 | 2 US-08-724-394A-20 | Sequence 20, Appl |
| 14 | 48.4 | 16.1 246240 | 2 US-08-724-394A-21 | Sequence 21, Appl |
| 15 | 48.4 | 16.1 246240 | 2 US-08-724-394A-22 | Sequence 22, Appl |
| 16 | 47.6 | 15.8 41684 | 4 US-09-536-059-1 | Sequence 1, Appl |
| 17 | 45.8 | 15.2 65042 | 4 US-09-784-316-3 | Sequence 3, Appl |
| 18 | 45 | 15.0 3757 | 4 US-09-620-312D-106 | Sequence 106, App |
| 19 | 43.8 | 14.6 392000 | 4 US-10-027-983-11 | Sequence 11, Appl |
| 20 | 42 | 14.0 786431 | 4 US-09-751-389-3 | Sequence 3, Appl |
| 21 | 41.6 | 13.8 168575 | 4 US-09-426-290-1 | Sequence 62, Appl |
| 22 | 39.8 | 13.2 6801 | 4 US-10-204-708-62 | Sequence 62, Appl |
| 23 | 38.8 | 12.9 55827 | 4 US-09-813-133A-3 | Sequence 3, Appl |
| 24 | 38.6 | 12.8 22067 | 4 US-09-820-001-3 | Sequence 3, Appl |
| 25 | 38.2 | 12.7 319608 | 4 US-09-539-33D-1 | Sequence 1, Appl |
| 26 | 38.2 | 12.7 319608 | 4 US-09-679-409-1 | Sequence 1, Appl |
| 27 | 38 | 12.6 15418 | 4 US-09-783-203-1 | Sequence 1, Appl |

| | | | | |
|----|------|-------------|-----------------------|-------------------|
| 28 | 38 | 12.6 51552 | 4 US-09-733-294A-30 | Sequence 30, Appl |
| 29 | 37.8 | 12.6 98844 | 4 US-09-791-211-10 | Sequence 10, Appl |
| 30 | 37.8 | 12.6 392000 | 4 US-10-027-983-11 | Sequence 11, Appl |
| 31 | 37.2 | 12.4 773 | 4 US-09-621-976-7972 | Sequence 7972, Ap |
| 32 | 36.6 | 12.2 73 | 4 US-09-621-976-8430 | Sequence 8430, Ap |
| 33 | 36.6 | 12.2 73 | 4 US-09-621-976-8766 | Sequence 8766, Ap |
| 34 | 36.6 | 12.2 73 | 4 US-09-621-976-8783 | Sequence 8783, Ap |
| 35 | 36.6 | 12.2 523 | 4 US-09-621-976-1442 | Sequence 1442, Ap |
| 36 | 36.6 | 12.2 788 | 4 US-09-288-143-27 | Sequence 27, Appl |
| 37 | 36.4 | 12.1 262 | 4 US-09-621-976-189 | Sequence 189, App |
| 38 | 36.2 | 12.0 6801 | 4 US-10-204-708-61 | Sequence 61, Appl |
| 39 | 35.8 | 11.9 116592 | 4 US-09-818-512-3 | Sequence 3, Appl |
| 40 | 35 | 11.6 2189 | 3 US-08-846-020A-1 | Sequence 1, Appl |
| 41 | 35 | 11.6 2189 | 4 US-09-617-871-1 | Sequence 1, Appl |
| 42 | 34.8 | 11.6 7218 | 1 US-08-232-463-14 | Sequence 14, Appl |
| 43 | 34.8 | 11.6 31208 | 4 US-09-852-067-3 | Sequence 3, Appl |
| 44 | 34.6 | 11.5 388 | 4 US-09-621-976-18393 | Sequence 18393, A |
| 45 | 34.6 | 11.5 427 | 4 US-09-621-976-18392 | Sequence 18392, A |

ALIGNMENTS

RESULT 1
US-09-801-876B-3
Sequence 3, Application US/09801876B
Patent No. 6492155
GENERAL INFORMATION:
APPLICANT: YE, Jane et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
FILE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: C1001160
CURRENT FILING DATE: 2001-03-09
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 148567
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(148567)
OTHER INFORMATION: n = A,T,C or G
US-09-801-876B-3

Query Match 20.1%; Score 60.4; DB 4; Length 148567;
Best Local Similarity 65.7%; Pred. No. 1,46-08;
Matches 88; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 16 TAGTAGAAGTGAAGACACACAGAAAGTGGAGAAACACTGCAATCATATAT 75
DB TATCAAAAAGTGAAGACTCCACAGAAATGGGAGAAATTTGCCAATATATAT 76328
QY 76 CTGATTAAGGTGTGATATATATATATATATATATATATATATATATATAT 135
DB CTACAAAGATGTCATTTCATATATCAAAAGGGGCGGCTGTGGCTGCTGTGT 76388
QY 136 TAAACCCCTTACC 149
DB 76389 AATCCACGCAATC 76402
RESULT 2
US-10-254-869-3
Sequence 3, Application US/10254869
Patent No. 6653117
GENERAL INFORMATION:
APPLICANT: YE, Jane et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
FILE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF

FILE REFERENCE: CL001160DIV
CURRENT APPLICATION NUMBER: US/10/254,869
CURRENT FILING DATE: 2002-09-26
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 148567
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(148567)
OTHER INFORMATION: n = A,T,C or G
US-10-254-869-3

Query Match 20.1%; Score 60.4; DB 4; Length 148567;
Best Local Similarity 65.7%; Pred. No. 1.4e-08;
Matches 88; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 16 TAGTAAAGAAAGTAAAGAAAGCAACCCAGAGAGTGGAGAAACACTTGCAATCATATAT 75
DB 76269 TATCAAAAAGTGAAGAAAGTCAACCCAGAGATGGAGAAATATTGCAAAATCATATAT 76328
QY 76 CTGATAGAGGTGTGATATATATATATATATAGTTTTCATAGTCTGCTGCTTA 135
DB 76329 CTACTAAAGATGTGATTTACATATATACAAAGGGCCAGCGCTGTGCTCATGCTGT 76388
QY 136 TAAACCCCTCACC 149
DB 76389 AATCCCGCAATC 76402

RESULT 3
US-09-804-471A-3/C
Sequence 3, Application US/09804471A
Patent No. 6479269
GENERAL INFORMATION:
APPLICANT: WEBSTER, Marion et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001164
CURRENT APPLICATION NUMBER: US/09/804,471A
CURRENT FILING DATE: 2001-03-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 174493
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(174493)
OTHER INFORMATION: n = A,T,C or G
US-09-804-471A-3

Query Match 18.5%; Score 55.8; DB 4; Length 174493;
Best Local Similarity 61.0%; Pred. No. 4.5e-07;
Matches 108; Conservative 0; Mismatches 67; Indels 2; Gaps 1;

QY 28 GAAAGCAACCCAGAGAGTGGAGAAACACTTGCAATCATATATCTGATAGGGTT 87
DB 69497 GAATGATACAGAAAGAGAGATGGATTAACAGAGACAGGTTGATATATATGA 69438
QY 88 GTGATATATGATATATATATATAGTTTTCATAGTCTGCTTATTAACCCCTCA 147
DB 69437 GATATATATATATATATATATAGTTTTCATCCAGGTTCTGCTTATATATATG 69378
QY 148 CCCTGTACAGTATTTGTTATATAGTTTGAAGTGGTGAAGCCGCAAGCAAAACT 204
DB 69377 TCCTGTACATCTTTTGTATTTAG--GGGTGCTTTAGGCTTTAGAACAGGCTT 69323

RESULT 4
US-10-238-709-3/C
Sequence 3, Application US/10238709
Patent No. 6680188
GENERAL INFORMATION:
APPLICANT: WEBSTER, Marion et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001164DIV
CURRENT APPLICATION NUMBER: US/10/238,709
CURRENT FILING DATE: 2002-09-11
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 174493
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(174493)
OTHER INFORMATION: n = A,T,C or G
US-10-238-709-3

Query Match 18.5%; Score 55.8; DB 4; Length 174493;
Best Local Similarity 61.0%; Pred. No. 4.5e-07;
Matches 108; Conservative 0; Mismatches 67; Indels 2; Gaps 1;

QY 28 GAAAGCAACCCAGAGAGTGGAGAAACACTTGCAATCATATATCTGATAGGGTT 87
DB 69497 GAATGATACAGAAAGAGAGATGGATTAACAGAGACAGGTTGATATATATGA 69438
QY 88 GTGATATATGATATATATATATAGTTTTCATAGTCTGCTTATTAACCCCTCA 147
DB 69437 GATATATATATATATATATATAGTTTTCATCCAGGTTCTGCTTATATATATG 69378
QY 148 CCCTGTACAGTATTTGTTATATAGTTTGAAGTGGTGAAGCCGCAAGCAAAACT 204
DB 69377 TCCTGTACATCTTTTGTATTTAG--GGGTGCTTTAGGCTTTAGAACAGGCTT 69323

RESULT 5
US-09-040-984-59/C
Sequence 59, Application US/09040984
Patent No. 6210883
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS
TITLE OF INVENTION: OF LUNG CANCER
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESSES:
ADDRESS: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,984
FILING DATE: 18-MAR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.456
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900

TELEFAX: 206-282-6031
 TELEX:
 INFORMATION FOR SEQ ID NO: 59:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 649 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-040-984-59

Query Match 17.7%; Score 53.2; DB 3; Length 649;
 Best Local Similarity 73.0%; Pred. No. 2.1e-07;
 Matches 81; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

QY 15 TTAGTAAAGGTTGTGATATATATATATATATAGTTTGTCCATAGT 125
 DB 312 TTATCGATTAAGGTGAAAACCAACCCACAGAA-TGGAGAAAATATTGGCAACCATGTA 254

QY 75 TCTGATTAAGGTTGTGATATATATATATATATAGTTTGTCCATAGT 125
 DB 253 TCTGATTAAGGTTGTGATATATATATATATATAGTTTGTCCATAGT 203

RESULT 6
 US-09-123-912-59/c
 Sequence 59, Application US/09123912A
 Patent No. 6312895
 GENERAL INFORMATION:
 APPLICANT: Reed, Steven G.
 APPLICANT: Wang, Tonglong
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
 FILE REFERENCE: 210121.455C1
 CURRENT APPLICATION NUMBER: US/09/123,912A
 CURRENT FILING DATE: 1998-07-27
 PRIOR APPLICATION NUMBER: 09/040,802
 PRIOR FILING DATE: 1998-03-18
 NUMBER OF SEQ ID NOS: 114
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 59
 LENGTH: 649
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (22)
 OTHER INFORMATION: Where n is a, c, g or t
 NAME/KEY: modified_base
 LOCATION: (190)
 OTHER INFORMATION: Where n is a, c, g or t
 NAME/KEY: modified_base
 LOCATION: (217)
 OTHER INFORMATION: Where n is a, c, g or t
 NAME/KEY: modified_base
 LOCATION: (430)
 OTHER INFORMATION: Where n is a, c, g or t
 NAME/KEY: modified_base
 LOCATION: (484)
 OTHER INFORMATION: Where n is a, c, g or t
 NAME/KEY: modified_base
 LOCATION: (544)
 OTHER INFORMATION: Where n is a, c, g or t
 NAME/KEY: modified_base
 LOCATION: (550)
 OTHER INFORMATION: Where n is a, c, g or t
 NAME/KEY: modified_base
 LOCATION: (577)
 OTHER INFORMATION: Where n is a, c, g or t
 NAME/KEY: modified_base
 LOCATION: (583)
 OTHER INFORMATION: Where n is a, c, g or t

NAME/KEY: modified_base
 LOCATION: (594)
 OTHER INFORMATION: Where n is a, c, g or t
 US-09-123-912-59

Query Match 17.7%; Score 53.2; DB 4; Length 649;
 Best Local Similarity 73.0%; Pred. No. 2.1e-07;
 Matches 81; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

QY 15 TTAGTAAAGGTTGTGATATATATATATATATAGTTTGTCCATAGT 125
 DB 312 TTATCGATTAAGGTGAAAACCAACCCACAGAA-TGGAGAAAATATTGGCAACCATGTA 254

QY 75 TCTGATTAAGGTTGTGATATATATATATATATAGTTTGTCCATAGT 125
 DB 253 TCTGATTAAGGTTGTGATATATATATATATATAGTTTGTCCATAGT 203

RESULT 7
 US-09-643-597-59/c
 Sequence 59, Application US/09643597
 Patent No. 6426072
 GENERAL INFORMATION:
 APPLICANT: Wang, Tonglong
 APPLICANT: Fan, Liqun
 APPLICANT: Kalos, Michael D.
 APPLICANT: Bangur, Chaitanya S.
 APPLICANT: Hosken, Nancy
 APPLICANT: Fanger, Gary R.
 APPLICANT: Li, Samuel X.
 APPLICANT: Wang, Aijun
 APPLICANT: Skeiky, Jasir A.W.
 APPLICANT: Henderson, Robert A.
 APPLICANT: McNeill, Patricia D.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 FILE REFERENCE: 210121.455C1
 CURRENT APPLICATION NUMBER: US/09/643,597
 CURRENT FILING DATE: 2000-08-21
 NUMBER OF SEQ ID NOS: 369
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 59
 LENGTH: 649
 TYPE: DNA
 ORGANISM: Homo sapien
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (1) -- (649)
 OTHER INFORMATION: n = A,T,C or G
 US-09-643-597-59

Query Match 17.7%; Score 53.2; DB 4; Length 649;
 Best Local Similarity 73.0%; Pred. No. 2.1e-07;
 Matches 81; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

QY 15 TTAGTAAAGGTTGTGATATATATATATATATAGTTTGTCCATAGT 125
 DB 312 TTATCGATTAAGGTGAAAACCAACCCACAGAA-TGGAGAAAATATTGGCAACCATGTA 254

QY 75 TCTGATTAAGGTTGTGATATATATATATATATAGTTTGTCCATAGT 125
 DB 253 TCTGATTAAGGTTGTGATATATATATATATATAGTTTGTCCATAGT 203

RESULT 8
 US-09-480-884A-59/c
 Sequence 59, Application US/09480884A
 Patent No. 6482597
 GENERAL INFORMATION:
 APPLICANT: Wang, Tonglong
 APPLICANT: Fan, Liqun
 APPLICANT: Hosken, Nancy A.
 APPLICANT: Kalos, Michael D.

APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
FILE REFERENCE: 210121.455C6
CURRENT APPLICATION NUMBER: US/09/480,884A
CURRENT FILING DATE: 2001-08-27
NUMBER OF SEQ ID NOS: 330
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 59
LENGTH: 649
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(649)
OTHER INFORMATION: n = A,T,C or G
US-09-480-884A-59

Query Match 17.7%; Score 53.2; DB 4; Length 649;
Best Local Similarity 73.0%; Pred. No. 2.1e-07;
Matches 81; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

15 TTAGTAAAGAGTAAAGCAACCCACAGAGTGGAGAGAAACCTTGCAATCATATA 74
312 TTATCGATTAAGTGAAGCAACCCACAGAA-TGGAGAGAAATATTGCAACCATGTA 254
75 TCTGATTAAGGTTGTGATATTATGATATATATATAGGTTTGTCCATAGT 125
253 TCTGATTAAGGTTGTGATATTCAGATCCAGATATATTAAGAAATCTTACACTCAGT 203

RESULT 9
US-09-542-615A-59/c
Sequence 59, Application US/09542615A
Patent No. 6518256
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hoeken, Nancy A.
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
FILE REFERENCE: 210121.455C8
CURRENT APPLICATION NUMBER: US/09/542,615A
CURRENT FILING DATE: 2000-04-14
NUMBER OF SEQ ID NOS: 350
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 59
LENGTH: 649
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(649)
OTHER INFORMATION: n = A,T,C or G
US-09-542-615A-59

Query Match 17.7%; Score 53.2; DB 4; Length 649;
Best Local Similarity 73.0%; Pred. No. 2.1e-07;
Matches 81; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

15 TTACTAAGAGTAAAGCAACCCACAGAGTGGAGAGAAACACTTGCAATCATATA 74
312 TTATCGATTAAGTGAAGCAACCCACAGAA-TGGAGAGAAATATTGCAACCATGTA 254
75 TCTGATTAAGGTTGTGATATTATGATATATATATAGGTTTGTCCATAGT 125
253 TCTGATTAAGGTTGTGATATTCAGATCCAGATATATTAAGAAATCTTACACTCAGT 203

RESULT 10

US-09-606-421B-59/c
Sequence 59, Application US/09606421B
Patent No. 6531315
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hoeken, Nancy
APPLICANT: Fanger, Gary R.
APPLICANT: Li, Samuel X.
APPLICANT: Skeiky, Yasir A.W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.455C9
CURRENT APPLICATION NUMBER: US/09/606,421B
CURRENT FILING DATE: 2000-06-28
NUMBER OF SEQ ID NOS: 358
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 59
LENGTH: 649
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(649)
OTHER INFORMATION: n = A,T,C or G
US-09-606-421B-59

Query Match 17.7%; Score 53.2; DB 4; Length 649;
Best Local Similarity 73.0%; Pred. No. 2.1e-07;
Matches 81; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

15 TTAGTAAAGAGTAAAGCAACCCACAGAGTGGAGAGAAACCTTGCAATCATATA 74
312 TTATCGATTAAGTGAAGCAACCCACAGAA-TGGAGAGAAATATTGCAACCATGTA 254
75 TCTGATTAAGGTTGTGATATTATGATATATATATAGGTTTGTCCATAGT 125
253 TCTGATTAAGGTTGTGATATTCAGATCCAGATATATTAAGAAATCTTACACTCAGT 203

RESULT 11
US-09-221-107-59/c
Sequence 59, Application US/09221107
Patent No. 6660838
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
FILE REFERENCE: 210121.455C2
CURRENT APPLICATION NUMBER: US/09/221,107
CURRENT FILING DATE: 1998-12-22
NUMBER OF SEQ ID NOS: 161
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 59
LENGTH: 649
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (22)
OTHER INFORMATION: Where n is a, c, g or t
FEATURE:
NAME/KEY: modified_base
LOCATION: (190)
OTHER INFORMATION: Where n is a, c, g or t
FEATURE:
NAME/KEY: modified_base
LOCATION: (217)
OTHER INFORMATION: Where n is a, c, g or t
FEATURE:
NAME/KEY: modified_base

15 TTACTAAGAGTAAAGCAACCCACAGAGTGGAGAGAAACACTTGCAATCATATA 74
312 TTATCGATTAAGTGAAGCAACCCACAGAA-TGGAGAGAAATATTGCAACCATGTA 254
75 TCTGATTAAGGTTGTGATATTATGATATATATATAGGTTTGTCCATAGT 125
253 TCTGATTAAGGTTGTGATATTCAGATCCAGATATATTAAGAAATCTTACACTCAGT 203

```
LOCATION: (430)
OTHER INFORMATION: where n is a, c, g or t
FEATURE:
NAME/KEY: modified_base
LOCATION: (433)
OTHER INFORMATION: where n is a, c, g or t
FEATURE:
NAME/KEY: modified_base
LOCATION: (484)
OTHER INFORMATION: where n is a, c, g or t
FEATURE:
NAME/KEY: modified_base
LOCATION: (544)
OTHER INFORMATION: where n is a, c, g or t
FEATURE:
NAME/KEY: modified_base
LOCATION: (550)
OTHER INFORMATION: where n is a, c, g or t
FEATURE:
NAME/KEY: modified_base
LOCATION: (577)
OTHER INFORMATION: where n is a, c, g or t
FEATURE:
NAME/KEY: modified_base
LOCATION: (583)
OTHER INFORMATION: where n is a, c, g or t
FEATURE:
NAME/KEY: modified_base
LOCATION: (594)
OTHER INFORMATION: where n is a, c, g or t
US-09-221-107-59

Query Match 17.7%; Score 53.2; DB 4; Length 649;
Best Local Similarity 73.0%; Pred. No. 2.1e-07;
Matches 81; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

Dy 15 TTACTAGAAAGTAAAGCAACCCACAGAGTGGAGAAAACCTTGCAATCATATA 74
312 TTATCGATAAAGTAAAGCAACCCACAGAGTGGAGAAAATATTGGCAACCATGTA 254

Qy 75 TCTGATAAGGTTGTGATATTATGATATATATATAGGTTTTTGTCCATAGT 125
253 TCTGATAAGGTTGTGATATTATGATATATATATAGGTTTTTGTCCATAGT 203

Db 253 TCTGATAAGGTTGTGATATTATGATATATATATAGGTTTTTGTCCATAGT 203

RESULT 12
US-09-016-434-1342
Sequence 1342, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
```

```
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 845-4166
TELEFAX: (650) 845-0555
INFORMATION FOR SEQ ID NO: 1342:
SEQUENCE CHARACTERISTICS:
LENGTH: 4080 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 934764
US-09-016-434-1342

Query Match 16.1%; Score 48.4; DB 4; Length 4080;
Best Local Similarity 70.9%; Pred. No. 1.7e-05;
Matches 78; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

Dy 14 CTTAGTAAGAAAGTAAAGCAACCCACAGAGTGGAGAAAACCTTGCAATCATAT 73
3159 CCTATCAAGAAAGTAAAGCAACCCACAGAGTGGAGAAAATTTTGGAGATCATAT 3217

Qy 74 ATCTGATAAGGTTGTGATATTATGATATATATATAGGTTTTTGTCCATA 123
3218 ATATGATAAGGTTGTGATATTATGATATATATATAGGTTTTTGTCCATA 3267

Db 3218 ATATGATAAGGTTGTGATATTATGATATATATATAGGTTTTTGTCCATA 3267

RESULT 13
US-08-724-394A-20/c
Sequence 20, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Krommel, Gregory S.
APPLICANT: Laufer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolfe, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
TITLE OF INVENTION: Sequences and Antibodies Thereto
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Pites, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
```

LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-20

Query Match 16.1%; Score 48.4; DB 2; Length 246240;
Best Local Similarity 77.8%; Pred. No. 0.00012;
Matches 84; Conservative 0; Mismatches 21; Indels 3; Gaps 2;

QY 186 AGGCTCAGAGCAAACTCTCTCTCAGCCCTCTCTCTCTGACACTTC 245
DB 173083 AGACCGCAGAGCCAAAGTTTCTCTGATCTCTCTGCGCCCTCTCTGAGTCC-C 173025
QY 246 ATTCTTCCCTGAGGCCA--CATAGAACTAGATCTCTTCCACAG 291
DB 173024 ATTCTCCCGAGCTAGCCATAGAAATGGAATTCCTTCTCAG 172977

RESULT 14
US-08-724-394A-21/C

/ Sequence 21, Application US/08724394A

/ Patent No. 5872237

/ GENERAL INFORMATION:

/ APPLICANT: Feder, John N.

/ APPLICANT: Krommal, Gregory S.

/ APPLICANT: Lauer, Peter M.

/ APPLICANT: Ruddy, David A.

/ APPLICANT: Thomas, Winston

/ APPLICANT: Teuchinashi, Zenta

/ APPLICANT: Wolff, Roger K.

/ TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1

/ TITLE OF INVENTION: Sequences and Antibodies Thereo

/ NUMBER OF SEQUENCES: 31

/ CORRESPONDENCE ADDRESSES:

/ ADDRESSES: TOWNSEND and TOWNSEND and CREM LLP

/ STREET: Two Embarcadero Center, 8th Floor

/ CITY: San Francisco

/ STATE: CA

/ COUNTRY: USA

/ ZIP: 94111-3834

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patentin Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/724,394A

/ FILING DATE: 01-OCT-1996

/ CLASSIFICATION: 536

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Fitts, Renee A.

/ REGISTRATION NUMBER: 35,136

/ REFERENCE/DOCKET NUMBER: 017957-000100

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: 415-576-0200

/ TELEFAX: 415-576-0300

/ INFORMATION FOR SEQ ID NO: 21:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 246240 base pairs

/ TYPE: nucleic acid

/ STRANDEDNESS: not relevant

/ TOPOLOGY: not relevant

/ MOLECULE TYPE: cDNA

/ FEATURE:

/ NAME/KEY: misc feature

/ LOCATION: 1..246240

/ OTHER INFORMATION: /note= "HLA-H.CONTIG"

US-08-724-394A-21

Query Match 16.1%; Score 48.4; DB 2; Length 246240;
Best Local Similarity 77.8%; Pred. No. 0.00012;
Matches 84; Conservative 0; Mismatches 21; Indels 3; Gaps 2;

QY 186 AGGCTCAGAGCAAACTCTCTCTCAGCCCTCTCTCTCTGACACTTC 245
DB 173083 AGACCGCAGAGCCAAAGTTTCTCTGATCTCTCTGCGCCCTCTCTGAGTCC-C 173025
QY 246 ATTCTTCCCTGAGGCCA--CATAGAACTAGATCTCTTCCACAG 291
DB 173024 ATTCTCCCGAGCTAGCCATAGAAATGGAATTCCTTCTCAG 172977

RESULT 15

US-08-724-394A-22/C

/ Sequence 22, Application US/08724394A

/ Patent No. 5872237

/ GENERAL INFORMATION:

/ APPLICANT: Feder, John N.

/ APPLICANT: Krommal, Gregory S.

/ APPLICANT: Lauer, Peter M.

/ APPLICANT: Ruddy, David A.

/ APPLICANT: Thomas, Winston

/ APPLICANT: Teuchinashi, Zenta

/ APPLICANT: Wolff, Roger K.

/ TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1

/ TITLE OF INVENTION: Sequences and Antibodies Thereo

/ NUMBER OF SEQUENCES: 31

/ CORRESPONDENCE ADDRESSES:

/ ADDRESSES: TOWNSEND and TOWNSEND and CREM LLP

/ STREET: Two Embarcadero Center, 8th Floor

/ CITY: San Francisco

/ STATE: CA

/ COUNTRY: USA

/ ZIP: 94111-3834

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patentin Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/724,394A

/ FILING DATE: 01-OCT-1996

/ CLASSIFICATION: 536

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Fitts, Renee A.

/ REGISTRATION NUMBER: 35,136

/ REFERENCE/DOCKET NUMBER: 017957-000100

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: 415-576-0200

/ TELEFAX: 415-576-0300

/ INFORMATION FOR SEQ ID NO: 22:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 246240 base pairs

/ TYPE: nucleic acid

/ STRANDEDNESS: not relevant

/ TOPOLOGY: not relevant

/ MOLECULE TYPE: cDNA

/ FEATURE:

/ NAME/KEY: misc feature

/ LOCATION: 1..246240

/ OTHER INFORMATION: /note= "HLA-H.CONTIG"

US-08-724-394A-22

Query Match 16.1%; Score 48.4; DB 2; Length 246240;
Best Local Similarity 77.8%; Pred. No. 0.00012;
Matches 84; Conservative 0; Mismatches 21; Indels 3; Gaps 2;

QY 186 AGGCTCAGAGCAAACTCTCTCTCAGCCCTCTCTCTCTGACACTTC 245
DB 173083 AGACCGCAGAGCCAAAGTTTCTCTGATCTCTCTGCGCCCTCTCTGAGTCC-C 173025

Oy 246 ATTCTTCCCTGAGGCA--CATAGAACTAGATCTCTCTTCCACAAG 291
Db 173024 ATTCTCCCGGAGTCTAGCCATAGAAATGAGAAATTCTCTCTCTCAAG 172977

Search completed: March 13, 2004, 00:50:58
Job time : 27.3958 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 21:22:51 / Search time 105.627 Seconds
(without alignments)
10491.276 Million cell updates/sec

Title: US-09-966-880a-10_COPY_6264_6564

Perfect score: 301
Sequence: 1 gacatcaagagacacttaga.....cttcacaaagcggtcgaag 301

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 243257 seqs, 184098884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 301 | 100.0 | 6564 | 9 | US-09-966-880a-10 |
| 2 | 77.6 | 25.8 | 511 | 15 | US-10-027-632-92073 |
| 3 | 77.6 | 25.8 | 511 | 15 | US-10-027-632-317886 |
| 4 | 74.4 | 24.7 | 3298 | 15 | US-10-027-632-112733 |
| 5 | 74.4 | 24.7 | 3298 | 15 | US-10-027-632-112734 |
| 6 | 74.4 | 24.7 | 164875 | 15 | US-10-085-117-332 |
| 7 | 73 | 24.3 | 367378 | 14 | US-10-312-841-1 |
| 8 | 67 | 22.3 | 35465 | 14 | US-10-161-572-6 |
| 9 | 67 | 22.3 | 35465 | 14 | US-10-161-572-6 |
| 10 | 66.6 | 22.1 | 568 | 15 | US-10-027-632-49688 |
| 11 | 66.6 | 22.1 | 1155 | 15 | US-10-027-632-117444 |
| 12 | 66.6 | 22.1 | 1770 | 15 | US-10-027-632-248967 |
| 13 | 66 | 21.9 | 431 | 12 | US-10-085-783A-24956 |
| 14 | 66 | 21.9 | 431 | 15 | US-10-242-535A-24956 |
| 15 | 65.8 | 21.9 | 1524 | 10 | US-09-764-891-5553 |

| | | | | | | |
|----|------|------|---------|----|----------------------|--------------------|
| 16 | 65.8 | 21.9 | 1524 | 14 | US-10-205-428-529 | Sequence 529, App |
| 17 | 65.6 | 21.8 | 521 | 15 | US-10-027-632-62410 | Sequence 62410, A |
| 18 | 65.6 | 21.8 | 521 | 15 | US-10-027-632-63014 | Sequence 63014, A |
| 19 | 65.6 | 21.8 | 521 | 15 | US-10-027-632-306536 | Sequence 306536, A |
| 20 | 65.2 | 21.7 | 593 | 15 | US-10-027-632-287770 | Sequence 287770, A |
| 21 | 65 | 21.6 | 452 | 10 | US-09-918-995-13639 | Sequence 13639, A |
| 22 | 64.4 | 21.4 | 13069 | 9 | US-09-764-869-1850 | Sequence 1850, Ap |
| 23 | 64.4 | 21.4 | 13069 | 14 | US-10-091-504-1850 | Sequence 1850, Ap |
| 24 | 64.4 | 21.4 | 13069 | 15 | US-10-027-577-1850 | Sequence 1850, Ap |
| 25 | 64.4 | 21.4 | 32203 | 9 | US-09-764-869-1849 | Sequence 1849, Ap |
| 26 | 64.4 | 21.4 | 32203 | 14 | US-10-091-504-1849 | Sequence 1849, Ap |
| 27 | 64.4 | 21.4 | 32203 | 15 | US-10-027-577-1849 | Sequence 1849, Ap |
| 28 | 62.6 | 20.8 | 1691139 | 14 | US-10-067-514-1 | Sequence 1, Appli |
| 29 | 62.6 | 20.8 | 1691139 | 15 | US-10-419-723-1 | Sequence 1, Appli |
| 30 | 62.4 | 20.7 | 531 | 14 | US-10-029-386-10489 | Sequence 10489, A |
| 31 | 62.4 | 20.7 | 1480 | 14 | US-10-017-161-337 | Sequence 337, App |
| 32 | 62.4 | 20.7 | 1480 | 15 | US-10-092-798-301 | Sequence 301, App |
| 33 | 62.4 | 20.4 | 315 | 12 | US-10-085-783A-30007 | Sequence 30007, A |
| 34 | 61.4 | 20.4 | 315 | 15 | US-10-242-535A-30007 | Sequence 30007, A |
| 35 | 60.4 | 20.1 | 148567 | 9 | US-09-801-876B-3 | Sequence 3, Appli |
| 36 | 60.4 | 20.1 | 148567 | 12 | US-10-667-442-3 | Sequence 3, Appli |
| 37 | 60.4 | 20.1 | 148567 | 14 | US-10-254-869-3 | Sequence 3, Appli |
| 38 | 59.4 | 19.7 | 604 | 15 | US-10-027-632-27507 | Sequence 27507, A |
| 39 | 59 | 19.6 | 604 | 15 | US-10-027-632-27506 | Sequence 27506, A |
| 40 | 58.2 | 19.3 | 39776 | 14 | US-10-160-293-3 | Sequence 3, Appli |
| 41 | 57.4 | 19.1 | 416 | 10 | US-09-918-995-36136 | Sequence 36136, A |
| 42 | 57.2 | 19.0 | 761 | 15 | US-10-027-632-15636 | Sequence 15636, A |
| 43 | 57 | 18.9 | 160361 | 12 | US-10-235-192A-35 | Sequence 35, Appli |
| 44 | 56.6 | 18.8 | 5879 | 14 | US-10-311-455-241 | Sequence 241, App |
| 45 | 56.4 | 18.7 | 3673778 | 14 | US-10-312-841-2 | Sequence 2, Appli |

ALIGNMENTS

| | |
|----------|--|
| RESULT 1 | US-09-966-880a-10 |
| 1 | Sequence 10, Application US/09966880A |
| 2 | Patent No. US2002016473A1 |
| 3 | GENERAL INFORMATION: |
| 4 | APPLICANT: Muramatsu, Masamichi |
| 5 | TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE |
| 6 | FILE REFERENCE: 06501-088001 |
| 7 | CURRENT APPLICATION NUMBER: US/09/966, 880A |
| 8 | CURRENT FILING DATE: 2001-09-28 |
| 9 | PRIOR APPLICATION NUMBER: PCT/JP00/01918 |
| 10 | PRIOR FILING DATE: 2000-03-28 |
| 11 | PRIOR APPLICATION NUMBER: JP 11-371382 |
| 12 | PRIOR FILING DATE: 1999-12-27 |
| 13 | PRIOR APPLICATION NUMBER: JP 11-178999 |
| 14 | PRIOR FILING DATE: 1999-06-24 |
| 15 | PRIOR APPLICATION NUMBER: JP 11-87192 |
| 16 | PRIOR FILING DATE: 1998-03-29 |
| 17 | NUMBER OF SEQ ID NOS: 36 |
| 18 | SOFTWARE: FASTSEQ for Windows Version 4.0 |
| 19 | SEQ ID NO 10 |
| 20 | LENGTH: 6564 |
| 21 | TYPE: DNA |
| 22 | ORGANISM: Homo sapiens |
| 23 | US-09-966-880a-10 |
| 24 | Query Match 100.0%; Score 301; DB 9; Length 6564; |
| 25 | Best Local Similarity 100.0%; Pred. No. 3.4e-81; Indels 0; Gaps 0; |
| 26 | Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| 27 | 1 GCATCAAGGACCTTAGTAGAAGTGAAGCAACCCACAGAGTGGAGAAAACAC 60 |
| 28 | 6264 GCATCAAGGACCTTAGTAGAAGTGAAGCAACCCACAGAGTGGAGAAAACAC 6263 |
| 29 | 61 TTGCAATCATATATCTGATTAAGGCTTGTGATATTAATATATATAGTTTGTCC 120 |
| 30 | 6324 TTGCAATCATATATCTGATTAAGGCTTGTGATATTAATATATATAGTTTGTCC 6383 |

| | | | |
|----|------|--|------|
| QY | 121 | ATAGTCTCGGCTTAATAACCCCTCACCCTGTAGACGATTTGTATAGGTTGAT | 180 |
| Db | 6384 | ATAGTCTCGGCTTAATAACCCCTCACCCTGTAGACGATTTGTATAGGTTGAT | 6443 |
| QY | 181 | GGTTTAGGCTCAGAGCAAACTCTCTCTCACTTCTCCAGCCCTCTGTCTTGGC | 240 |
| Db | 6444 | GGTTTAGGCTCAGAGCAAACTCTCTCTCACTTCTCCAGCCCTCTGTCTTGGC | 6503 |
| QY | 241 | ACCTCATTTCTTCCCTGAGGCCACATGAAACTAGAACTCTTTCACAAAGGCGTCAA | 300 |
| Db | 6504 | ACCTCATTTCTTCCCTGAGGCCACATGAAACTAGAACTCTTTCACAAAGGCGTCAA | 6563 |
| QY | 301 | G 301 | |
| Db | 6564 | G 6564 | |

RESULT 2

```

? Sequence: 92073, Application US/10027632
? Publication No. US20030204075A9
? GENERAL INFORMATION:
?   APPLICANT: Wang, David G.
?   TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
?     TITLE OF INVENTION: Polymorphisms in the Human Genome
?   FILE REFERENCE: 108827.129
?   CURRENT APPLICATION NUMBER: US/10/027.632
?   CURRENT FILING DATE: 2002-04-30
?   PRIOR APPLICATION NUMBER: US 60/218, 006
?   PRIOR FILING DATE: 2000-07-12
?   PRIOR APPLICATION NUMBER: US 60/198, 676
?   PRIOR FILING DATE: 2000-04-20
?   PRIOR APPLICATION NUMBER: US 60/193, 483
?   PRIOR FILING DATE: 2000-03-29
?   PRIOR APPLICATION NUMBER: US 60/185, 218
?   PRIOR FILING DATE: 2000-02-24
?   PRIOR APPLICATION NUMBER: US 60/167, 363
?   PRIOR FILING DATE: 1999-11-23
?   PRIOR APPLICATION NUMBER: US 60/156, 358
?   PRIOR FILING DATE: 1999-09-28
?   PRIOR APPLICATION NUMBER: US 60/146, 002
?   PRIOR FILING DATE: 1999-08-09
?   NUMBER OF SEQ ID NOS: 325720
?   SOFTWARE: FastSeq for Windows Version 4.0
?   SEQ ID NO 92073
?   LENGTH: 511
?   TYPE: DNA
?   ORGANISM: Human
? US-10-027-632-92073

```

| | | | | |
|-----------------------|--------------|--------------------|----------------|-------------|
| Query Match | 25.8%; | Score 77.6; | DB 15; | Length 511; |
| Best Local Similarity | 68.6%; | Pred. No. 2.2e-13; | | |
| Matches 107; | Conservative | 0; | Mismatches 49; | Indels 0; |
| | | | Gaps | 0 |

| | | | |
|----|-----|--|------|
| QY | 66 | AATCATATATCTGATGAAGGTTGTGATATATATATATATATATAGTGTTCCTCATGT | 125 |
| Db | 463 | ACTATTATTAAGCTTACTTTGGCAATGTGGTGTATGATATACATGGTATTGTTGTCCACAGT | 4040 |
| QY | 126 | TCCTGGCTTATTAACCCCTCACCCCTGTGTACAGTCAATTGTTATATAGTTGGATGGTTT | 185 |
| Db | 403 | TCTCGGATCATTAATCCCATAGCCCTGTGTACAGCTCTTTTCATATATGTTGGGTGTGT | 344 |
| QY | 186 | AGGCTCAGAAGCAAACTCTCTCTCTCAACCTTCTC | 221 |
| Db | 343 | AGGCTTCAGGACAGGCTCTTACTTCTCTCTCCCTC | 308 |

RESULT 3

US-10-027-632-317886/c
; Sequence 317886, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:

```

1  APPLICANT: Wang, David G.
2  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
3  TITLE OF INVENTION: Polymorphisms in the Human Genome
4  FILE REFERENCE: 106827.129
5  CURRENT APPLICATION NUMBER: US/10/027,632
6  CURRENT FILING DATE: 2002-04-30
7  PRIOR APPLICATION NUMBER: US 60/218,006
8  PRIOR FILING DATE: 2000-07-12
9  PRIOR APPLICATION NUMBER: US 60/198,676
10 PRIOR FILING DATE: 2000-04-20
11 PRIOR APPLICATION NUMBER: US 60/193,483
12 PRIOR FILING DATE: 2000-03-29
13 PRIOR APPLICATION NUMBER: US 60/185,218
14 PRIOR FILING DATE: 2000-02-24
15 PRIOR APPLICATION NUMBER: US 60/167,363
16 PRIOR FILING DATE: 1999-11-23
17 PRIOR APPLICATION NUMBER: US 60/156,358
18 PRIOR FILING DATE: 1999-09-28
19 PRIOR APPLICATION NUMBER: US 60/146,002
20 PRIOR FILING DATE: 1998-08-09
21 NUMBER OF SEQ ID NOS: 325720
22 SOFTWARE: FastSeq for Windows Version 4.0
23 SEQ ID NO 317886
24 LENGTH: 511
25 TYPE: DNA
26 ORGANISM: Human
27 US-10-027-632-317886

```

| | | | | |
|-----------------------|-----------------|--------------------|-----------|-------------|
| Query Match | 25.8% | Score 77.6; | DB 15; | Length 511; |
| Best Local Similarity | 68.6%; | Pred. No. 2.2e-13; | | |
| Matches 107; | Conservative 0; | Mismatches 49; | Indels 0; | Gaps 0; |

| | | | |
|----|-----|--|-----|
| Qy | 66 | AATCATATATCTGATAAAGGTTGTGATATATATGAAATATATAGTTTTCACATAGT | 129 |
| Db | 463 | ACTTATTAAGCTTACTTTGGCAGTGTGGTGTATGATGATTAACATTTGGTTTGTCCACAGT | 404 |
| Qy | 126 | TCTTGAGCTTATTAACCCCTCACCCTGTGTTACAGTCATTGTGTTAAAGTTGATGGTTT | 185 |
| Db | 403 | TCTTGATCATTAACCTCCCATACCCCTGTGTTACAGTCTTTTCTCATTAATGTTGGTGTGTT | 344 |
| Qy | 186 | AGGCTCTAGAGCAAAATCTCTCTCTACCTTC | 221 |
| Db | 343 | AGGCTCTAGAGCAAGCCTCTTACTTTTCTCTGCCC | 308 |

RESULT 4

```

US-10-027-632-112733/c
/ Sequence 112733, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ POLYMORPHISMS IN THE HUMAN GENOME
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ. ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 112733

```


| Accession | Sequence | Position |
|-----------|--|----------|
| Db | ATTTCACACCTTAAATAATCTATACGTATTAATTTATATATAAATTTATACCATAT | 1093210 |
| Qy | TCCTGCGCTTATAAACCCCTCACCCCTGTTCAGTCATTTGTATTAAGTTGAGTGT | 126 |
| Db | TCCCTACCTTATTAATTCCTCATACCTTATTCACATCTATTTATTAATTAATTAAC | 1093150 |
| Qy | AGGCTCTAGAAGCAAAACTC-----TCCTCTCAGCTCTCTCAGCGCCCTCTC | 186 |
| Db | AAACCTCAAAAACAATCTCAATAAAGCAAAATCTCTCTAACCCTTCTCTACCCCTCAT | 1093090 |
| Qy | TCCTGGACCTCATTTCTTCCCTGAGGCGCATAGAAACTGATCATCTCTCT | 235 |
| Db | ATCTACTCTTATCTCTCCAAAACAAAATCTTTACCATCTTTCTAT | 1093030 |

```

RESULT 8
US-10-161-572-6/c
; Sequence 6, Application US/10161572
; Publication No. US20030087266A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: IGS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-097C-PC
; CURRENT APPLICATION NUMBER: US/10/161,572
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 35465
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-161-572-6

```

| | | | | | | |
|----|-----------------------|---|-------------------|----------|-------------|--|
| | Query Match | 22.3% | Score 67 | DB 14 | Length 3545 | |
| | Best Local Similarity | 72.7% | Pred. No. 3,6e-09 | | | |
| | Matches 113 | Conservative 0 | Mismatches 45 | Indels 9 | Gaps 1 | |
| QY | 85 | GTGTGATATATGATATATATATAGGTTTTGTGCATAGTTCCTGGCTTAATAACCC | 144 | | | |
| DB | 4807 | GTGGATATATATATATATATATATGTTTCATCCACAGTTCCTGGCTGTAATCCTA | 4748 | | | |
| QY | 145 | TCACCCCTGTTACAGTCATTTGTTATAAGGTTGATGGTTTAGGCC-----TCAGA | 195 | | | |
| DB | 4747 | CAGCCTTTGTTACAGTCCTCTGTTTAAATGTTGTATACCTTAAGGCTGAGGACAGGCTCGA | 4688 | | | |
| QY | 136 | AGCAAAACTCTCTCTCTACCTTTCCAGCCCTCCTGCTCTGGAC | 242 | | | |
| DB | 4687 | AGAAACAAATCTCTCTGACCTCTCTCGCCCTCCATTCCTGCCCC | 4641 | | | |

```

RESULT 9
US-10-161-572-8/C
; Sequence 8, Application US/10161572
; Publication No. US2003087266A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: IGS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-097C-PC
; CURRENT APPLICATION NUMBER: US/10/161,572
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605

```

```

; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,723
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 36891
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-161-572-8

```

| | | | | |
|-----------------------|----------------|-------------------|----------|--------------|
| Query March | 22.3% | Score 67 | DB 14 | Length 36991 |
| Best Local Similarity | 67.7% | Pred. No. 3,7e-09 | | |
| Matches 113 | Conservative 0 | Mismatches 45 | Indels 9 | Gaps 1 |

| | | | |
|----|------|---|------|
| QY | 85 | GTGTGATATATGATATATATATAGGTTTTGTCATAGTCTGTGCCTATTAACCC | 144 |
| DB | 4807 | GTGTATATTATATATATATATATATATGTTTCATCCACAGTCTGTGCTGTAAC | 4748 |
| QY | 145 | TCACCCCTTTACAGCTATTGTTATAGAAGTTGATAGTTTAGCC-----TCGA | 195 |
| DB | 4747 | CAGCCTTTGTAACAGTCTCTGTTATATATGTTGTAACCTTAGAGCTGGGGCAGGTGA | 4688 |
| QY | 196 | AGCAAAATCTCTCTTCACCTTTCCAGCCCTCTGTCTGTGCAC | 242 |
| DB | 4687 | AGAAACGAATCTCTTCAGCCTCTCTCTGCCCCTCATTTCTGCCCC | 4661 |

```

/ RESULT 10
/ US-10-027-632-49688/c
/ Sequence 49688, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Mann, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: Polymorphisms in the Human Genome
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 49688
/ LENGTH: 568
/ TYPE: DNA
/ ORGANISM: Human
/ US-10-027-632-49688

```

| | | | | | | | |
|-----------------------|-------|---|-------|------------|----|--------|-----|
| Query Match | 22.1% | Score | 66.6 | DB | 15 | Length | 568 |
| Best Local Similarity | 67.9% | Pred. | No. 5 | 5e-10 | | | |
| Matches | 93 | Conservative | 0 | Mismatches | 44 | Indels | 0 |
| | | | | Gaps | 0 | | |
| Qy | 85 | GTGTGATATTATGATATATATATATATGTTTTTGTGCCAATGTTCCGTGCTTATTAACCCC | 144 | | | | |
| Db | 332 | GTGCGTATATTATGATATATATTTGGTTTTCATCATGAGTTCCGSGCTCAAACTCGA | 263 | | | | |

| QY | Db | QY | Db |
|-----|--|-----|----|
| 145 | TCACCCCTGTTAAGCATTTGTATTAAGGTGATTTAGGCTCAGAAGCAAACT | 204 | |
| 262 | TAGCCCTCTTAAGCTTTTGTCAATACTTGGGTGCTTAGGCTCAGGGGAGAGGCT | 203 | |
| 205 | CTCTCTCTACCTTTC | 221 | |
| 202 | CTGACCTTCTCTGCCC | 186 | |

```

RESULT 11
US-10-027-632-117444
; Sequence 117444, Application US/10027632
; Publication No. US20030204075A3
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; RIF. REFERENCES: 108927,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ. ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 117444
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-117444

```

| | Query Match | 22.1% | Score 66.6 | DB 15 | Length 1155 |
|----|-----------------------|---|-----------------|---------------|-------------|
| | Best Local Similarity | 71.9% | Pred. No. 8e-10 | Mismatches 34 | Indels 0 |
| | Matches | 87 | Conservative | 0 | Gaps 0 |
| Oy | 99 | ATAATATATAGGTTTTCATGATGTTCTCGCTTAATAAACCCCTCACCTGTGTACA | 158 | | |
| Db | 784 | ATGATATGATTCGCTTTTCATCCAGCTTCCTGCGTCATPACTGCATATGTCCTGATGCA | 843 | | |
| Oy | 159 | GTCATTGTTATAAAGGTTGATGCTTTAGGCTCAGAAACAAACTCTCTCTACCTT | 218 | | |
| Db | 844 | AACCTTTGTTATATATGTTGATGCTTTAGGCTCAGGAAACAGATCTCTCTCGCCTT | 903 | | |
| Oy | 219 | C | 219 | | |
| Db | 904 | C | 904 | | |

```

RESULT 12
US-10-027-632-248967/c
; Sequence 248967, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676

```

```

? PRIOR FILING DATE: 2000-04-20
? PRIOR APPLICATION NUMBER: US 60/193,483
? PRIOR FILING DATE: 2000-03-29
? PRIOR APPLICATION NUMBER: US 60/185,218
? PRIOR FILING DATE: 2000-08-24
? PRIOR APPLICATION NUMBER: US 60/167,363
? PRIOR FILING DATE: 1999-11-23
? PRIOR APPLICATION NUMBER: US 60/156,358
? PRIOR FILING DATE: 1999-03-28
? PRIOR APPLICATION NUMBER: US 60/146,002
? PRIOR FILING DATE: 1999-08-09
? NUMBER OF SEQ ID NOS: 325720
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO: 248967
? LENGTH: 1770
? TYPE: DNA
? ORGANISM: Human
US-10-027-632-248967

```

| | | | | |
|-----------------------|--------------|-----------------|---------------|-------------|
| Query Match | 22.1% | Score 66.6 | DB 15 | Length 1770 |
| Best Local Similarity | 71.9% | Pred. No. 1e-09 | | |
| Matches 87 | Conservative | 0 | Mismatches 34 | Indels 0 |
| | | | Gaps | 0 |

| | | | |
|----|-----|--|-----|
| Oy | 99 | AATATAATPAAAGTTTGTGCAATGTCCTGGCTATTAACCCCGCACCGTTGACA | 158 |
| | | | |
| D6 | 982 | ATAGAATGAAGTAGTGTTTCATCCACAGTTCCTGGCTCATTAAGTCGCAATAGTCTTAGTGCA | 923 |
| Oy | 159 | GTCATTGTTAATPAAGTTGGATGATTTAAGGCCTCAGAAAGCAAACCTCTCTGCACCTT | 218 |
| | | | |
| D6 | 922 | AAC TT T G T A T P A A G T T G A T G T G T T A G G C C T A G A A A C A G A T C T C T C C T G C C T T | 863 |
| Oy | 219 | C 219 | |
| | | | |
| D6 | 862 | C 862 | |

RESULT 13
 US-10-085-783A-24956/C
 Sequence 24956, Application US/10085783A
 Publication No. US20040037841A1
 GENERAL INFORMATION:
 APPLICANT: ChondroGene Inc.
 APPLICANT: Liew, C.C.
 TITLE OR INVENTION: Compositions and Methods Relating to Osteoarthritis
 FILE REFERENCE: 4231/2002
 CURRENT APPLICATION NUMBER: US/10/085,783A
 CURRENT FILING DATE: 2002-02-28
 PRIOR APPLICATION NUMBER: US 60/305,340
 PRIOR FILING DATE: 2001-07-13
 PRIOR APPLICATION NUMBER: US 60/275,017
 PRIOR FILING DATE: 2001-03-12
 PRIOR APPLICATION NUMBER: US 60/271,955
 PRIOR FILING DATE: 2001-02-28
 NUMBER OF SEQ. ID NOS: 58994
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 24956

```

? TYPE: DNA
? ORGANISM: Human
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (3)..(3)
? OTHER INFORMATION: n is a, c, g, or t
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (5)..(5)
? OTHER INFORMATION: n is a, c, g, or t
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (29)..(29)
? OTHER INFORMATION: n is a, c, g, or t
? FEATURE:
? NAME/KEY: misc_feature

```

Search completed: March 13, 2004, 05:30:45
Job time : 116.627 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 12:27:59 ; Search time 273.899 Seconds
(without alignments)
137677.261 Million cell updates/sec

Title: US-09-966-880a-11
Perfect score: 87
Sequence: 1 agagaccatcatatcga.....ctgacaccactatgacag 87

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GeneBdb1:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: gb_vt:*
16: em_ba:*
17: em_fun:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_vt:*
31: em_htg_hum:*
32: em_htg_inv:*
33: em_htg_other:*
34: em_htg_mus:*
35: em_htg_pln:*
36: em_htg_rnd:*
37: em_htg_rnd:*
38: em_htg_vrt:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|---------------------|
| 1 | 87 | 100.0 | 87 | BD016836 | BD016836 Novel cyt |
| 2 | 87 | 100.0 | 2818 | BD016833 | BD016833 Novel cyt |
| 3 | 87 | 100.0 | 5514 | BD016834 | BD016834 Novel cyt |
| 4 | 87 | 100.0 | 11204 | BD016860 | BD016860 Novel cyt |
| 5 | 87 | 100.0 | 11204 | AB040430 | AB040430 Homo sapi |
| 6 | 87 | 100.0 | 71132 | AC092184 | AC092184 Homo sapi |
| 7 | 84 | 96.6 | 1828 | BC006296 | BC006296 Homo sapi |
| 8 | 84 | 96.6 | 2791 | AB040431 | AB040431 Homo sapi |
| 9 | 41 | 47.1 | 1767 | AB091291 | AB091291 Mus muscu |
| 10 | 41 | 47.1 | 2440 | BD016828 | BD016828 Novel cyt |
| 11 | 41 | 47.1 | 2440 | AF132979 | AF132979 Mus muscu |
| 12 | 41 | 47.1 | 178130 | AC119975 | AC119975 Mus muscu |
| 13 | 34.6 | 39.8 | 241757 | AC094826 | AC094826 Rattus no |
| 14 | 34.6 | 39.8 | 259506 | AC109119 | AC109119 Rattus no |
| 15 | 34.6 | 39.8 | 345098 | AC120617 | AC120617 Rattus no |
| 16 | 32.6 | 37.5 | 228932 | AL645669 | AL645669 Mouse DNA |
| 17 | 32.6 | 37.5 | 231634 | AC101661 | AC101661 Mus muscu |
| 18 | 32.2 | 37.0 | 175970 | AC097488 | AC097488 Homo sapi |
| 19 | 31.4 | 36.1 | 63101 | AP003489 | AP003489 Oryza sat |
| 20 | 31.4 | 36.1 | 157739 | AC122485 | AC122485 Mus muscu |
| 21 | 31.4 | 36.1 | 160545 | AP003508 | AP003508 Oryza sat |
| 22 | 31.2 | 35.9 | 165953 | AC129243 | AC129243 Human DNA |
| 23 | 31.2 | 35.9 | 288358 | AC106121 | AC106121 Rattus no |
| 24 | 31.2 | 35.6 | 137520 | EC4459965 | EC4459965 Equus cab |
| 25 | 30.6 | 35.2 | 157735 | CNS01RGY | AL160234 Human chr |
| 26 | 30.6 | 34.5 | 265272 | AC132111 | AC132111 Mus muscu |
| 27 | 29.8 | 34.3 | 184037 | AC146658 | AC146658 Oryza sat |
| 28 | 29.8 | 34.3 | 193410 | AC020594 | AC020594 Homo sapi |
| 29 | 29.6 | 34.0 | 179124 | AL603710 | AL603710 Mouse DNA |
| 30 | 29.6 | 34.0 | 237087 | AC128203 | AC128203 Rattus no |
| 31 | 29.6 | 34.0 | 247899 | AC025909 | AC025909 Mus muscu |
| 32 | 29.6 | 34.0 | 247899 | AC108233 | AC108233 Rattus no |
| 33 | 29.6 | 34.0 | 249713 | AC119873 | AC119873 Mus muscu |
| 34 | 29.6 | 34.0 | 260894 | AC122982 | AC122982 Rattus no |
| 35 | 29.6 | 34.0 | 261378 | AC079505 | AC079505 Mus muscu |
| 36 | 29.4 | 33.8 | 185694 | AC079528 | AC079528 Mus muscu |
| 37 | 29.4 | 33.8 | 213674 | AC098443 | AC098443 Rattus no |
| 38 | 29.4 | 33.6 | 123325 | AP006116 | AP006116 Lotus cor |
| 39 | 29.2 | 33.6 | 152422 | AC016491 | AC016491 Homo sapi |
| 40 | 29.2 | 33.6 | 164519 | AC020630 | AC020630 Homo sapi |
| 41 | 29.2 | 33.6 | 215353 | AC137981 | AC137981 Mus muscu |
| 42 | 29.2 | 33.6 | 231074 | AC142509 | AC142509 Rattus no |
| 43 | 29.2 | 33.6 | 231140 | AC097799 | AC097799 Rattus no |
| 44 | 29.2 | 33.3 | 76836 | AC119618 | AC119618 Homo sapi |
| 45 | 29 | | | | |

ALIGNMENTS

RESULT 1
BD016836
LOCUS BD016836
DEFINITION Novel cytidine deaminase.
ACCESSION BD016836
VERSION BD016836.1 GI:22558012
KEYWORDS JP 2001245669-A/9.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eulalia; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 87)
AUTHORS Honjo, T. and Muramatsu, M.
TITLE Novel cytidine deaminase
JOURNAL Patent: JP 2001245669-A 9 11-SEP-2001;

COMMENT JAPAN TOBACCO INC,TASUKU HONJO
OS Homo sapiens (human)
PN JP 2001245669-A/9
PD 11-SEP-2001
PF 28-MAR-2000 JP 2000092981
PI TASUKU HONJO,MASAMICHI MURAMATSU
PC C12N15/09,A61K39/395,A61K39/395,A61P1/00,A61P11/06,A61P13/12,
PC A61P17/00,
PC A61P27/02,A61P37/02,A61P37/08,C07K16/18,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12N9/78,C12P21/02,C12P21/08/(C12N1/21,C12R1:19), PC
(C12N5/10,C12R1:91),C12N15/00,C12N5/00,(C12N5/00,C12R1:91) CC
FH Key Location/Qualifiers
1..87
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
1..87
Location/Qualifiers

ORIGIN
Query Match 100.0%; Score 87; DB 6; Length 87;
Best Local Similarity 100.0%; Pred. No. 1e-20;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGACCATCAATTAATGAGTATTTCTGGCCCTGAGACTTGAGGGAGGCAAGA 60
DB 1 AGAAGACCATCAATTAATGAGTATTTCTGGCCCTGAGACTTGAGGGAGGCAAGA 60

QY 61 AGACACTCTGGACACCACTATGAGCAG 87
DB 61 AGACACTCTGGACACCACTATGAGCAG 87

RESULT 2
BD016833
LOCUS BD016833 2818 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel cytidine deaminase.
ACCESSION BD016833.1 GI:22558009
VERSION JP 2001245669-A/6.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2818)
AUTHORS Honjo,T. and Muramatsu,M.
TITLE Novel cytidine deaminase
JOURNAL Patent: JP 2001245669-A 6 11-SEP-2001;
JAPAN TOBACCO INC,TASUKU HONJO
COMMENT OS Homo sapiens (human)
PN JP 2001245669-A/6
PD 11-SEP-2001
PF 28-MAR-2000 JP 2000092981
PI TASUKU HONJO,MASAMICHI MURAMATSU
PC C12N15/09,A61K39/395,A61K39/395,A61P1/00,A61P11/06,A61P13/12,
PC A61P17/00,
PC A61P27/02,A61P37/02,A61P37/08,C07K16/18,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12N9/78,C12P21/02,C12P21/08/(C12N1/21,C12R1:19), PC
(C12N5/10,C12R1:91),C12N15/00,C12N5/00,(C12N5/00,C12R1:91) CC
FH Key Location/Qualifiers
1..87
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
1..2818
Location/Qualifiers

ORIGIN
Query Match 100.0%; Score 87; DB 6; Length 2818;
Best Local Similarity 100.0%; Pred. No. 1.2e-20;

Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGACCATCAATTAATGAGTATTTCTGGCCCTGAGACTTGAGGGAGGCAAGA 60
DB 1 AGAAGACCATCAATTAATGAGTATTTCTGGCCCTGAGACTTGAGGGAGGCAAGA 60

QY 61 AGACACTCTGGACACCACTATGAGCAG 87
DB 61 AGACACTCTGGACACCACTATGAGCAG 87

RESULT 3
BD016834
LOCUS BD016834 5514 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel cytidine deaminase.
ACCESSION BD016834.1 GI:22558010
VERSION JP 2001245669-A/7.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 5514)
AUTHORS Honjo,T. and Muramatsu,M.
TITLE Novel cytidine deaminase
JOURNAL Patent: JP 2001245669-A 7 11-SEP-2001;
JAPAN TOBACCO INC,TASUKU HONJO
COMMENT OS Homo sapiens (human)
PN JP 2001245669-A/7
PD 11-SEP-2001
PF 28-MAR-2000 JP 2000092981
PI TASUKU HONJO,MASAMICHI MURAMATSU
PC C12N15/09,A61K39/395,A61K39/395,A61P1/00,A61P11/06,A61P13/12,
PC A61P17/00,
PC A61P27/02,A61P37/02,A61P37/08,C07K16/18,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12N9/78,C12P21/02,C12P21/08/(C12N1/21,C12R1:19), PC
(C12N5/10,C12R1:91),C12N15/00,C12N5/00,(C12N5/00,C12R1:91) CC
FH Key Location/Qualifiers
1..5514
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
1..5514
Location/Qualifiers

ORIGIN
Query Match 100.0%; Score 87; DB 6; Length 5514;
Best Local Similarity 100.0%; Pred. No. 1.3e-20;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGACCATCAATTAATGAGTATTTCTGGCCCTGAGACTTGAGGGAGGCAAGA 60
DB 1032 AGAAGACCATCAATTAATGAGTATTTCTGGCCCTGAGACTTGAGGGAGGCAAGA 1091

QY 61 AGACACTCTGGACACCACTATGAGCAG 87
DB 1092 AGACACTCTGGACACCACTATGAGCAG 1118

RESULT 4
BD016860
LOCUS BD016860 11204 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel cytidine deaminase.
ACCESSION BD016860.1 GI:22558036
VERSION JP 2001245669-A/33.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 11204)
 AUTHORS Honjo, T. and Muramatsu, M.
 TITLE Novel cytidine deaminase
 JOURNAL Patent: JP 2001245668-A 33 11-SEP-2001;
 JAPAN TOBACCO INC, TASUKU HONJO
 COMMENT OS Homo sapiens (human)
 PN JP 2001245669-A/33
 PD 11-SEP-2001
 PF 26-MAR-2000 JP 2000092981
 PI TASUKU HONJO, MASAMICHI MORIMATSU
 PC C12N15/09, A61K39/395, A61P1/00, A61P11/06, A61P13/12,
 PC A61P17/00,
 PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC
 C12N1/21,
 PC C12N5/10, C12N9/78, C12P21/02, C12P21/08// (C12N1/21, C12R1.19), PC
 (C12N5/10, C12R1.91), C12N5/00, C12N5/00, (C12N5/00, C12R1.91) CC
 FH Key
 FEATURES Location/Qualifiers
 source 1..11204
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

Query Match 100.0%; Score 87; DB 6; Length 11204;
 Best Local Similarity 100.0%; Pred. No. 1.3e-20;
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGACATCATTAATTAAGTGAATTTTCTGGCCCTGAGACTTGACGGAGGCAAGA 60
 DB 442 AGAAGACATCATTAATTAAGTGAATTTTCTGGCCCTGAGACTTGACGGAGGCAAGA 501

QY 61 AGAAGCTCTGGACACCATATGACAG 87
 DB 502 AGAAGCTCTGGACACCATATGACAG 528

RESULT 5
 AB040430 11204 bp DNA linear PRI 03-OCT-2000
 LOCUS Homo sapiens AID gene for activation-induced cytidine deaminase,
 DEFINITION complete cds.
 ACCESSION AB040430
 VERSION AB040430.1 GI:9988407
 KEYWORDS AID; activation-induced cytidine deaminase.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (sites)
 Muto, T., Muramatsu, M., Tanigaki, M., Kinoshita, K. and Honjo, T.
 Isolation, tissue distribution, and chromosomal localization of the
 human activation-induced cytidine deaminase (AID) gene
 Genomics 68 (1), 85-88 (2000)

JOURNAL MEDLINE 20408890
 PUBMED 10950930
 REFERENCE 2 (sites)
 Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O.,
 Catalan, N., Forveille, M., Dufourcq-Lageouse, R., Gennery, A.,
 Tetzan, I., Ercsey, F., Kayserili, H., Ugazio, A. G., Brousse, N.,
 Muramatsu, M., Notarangelo, L. D., Kinoshita, K., Honjo, T., Fischer, A.
 and Durandy, A.
 Activation-induced cytidine deaminase (AID) deficiency causes the
 autosomal recessive form of the Hyper-IgM syndrome (HIGM2)
 Cell 102 (5), 565-575 (2000)

JOURNAL MEDLINE 20460541
 PUBMED 11007475
 REFERENCE 3 (bases 1 to 11204)
 Muto, T., Muramatsu, M., Tanigaki, M., Kinoshita, K. and Honjo, T.
 Direct Substitution
 Submitted (18-MAR-2000) Tasuku Honjo, Kyoto University, Department
 of Medical Chemistry, Faculty of Medicine, Yoshida, Sakyo-ku,
 Kyoto, Kyoto 606-8501, Japan (E-mail: honjo@four.med.kyoto-u.ac.jp).

FEATURES
 source
 Location/Qualifiers
 1..11204
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 join(521..528,6280..6427,7807..8077,8371..8486,8956..9009)
 /gene="AID"
 join(521..528,6280..6427,7807..8077,8371..8486,8956..9009)
 /gene="AID"
 /codon_start=1
 /product="activation-induced cytidine deaminase"
 /protein_id="BAB12720.1"
 /db_xref="GI:9988408"
 /translation="MDSLMMNRKFIYGRKNRMAGRETYCYVVRKRDSTFSL
 DFGYLRKNGCHVELFLRYISDMWDHPGRCVRIWFTSMSPCYDCAHVAHDFRGNP
 NLSIRIFTNALYPCEDRKAEPEGLRRLHRAQVIAIMTKDYFCWNTVEHNERTEK
 AMEGLHNSVRLSRQLRILLPLVEVDIDRDAFRITGL"

Query Match 100.0%; Score 87; DB 6; Length 11204;
 Best Local Similarity 100.0%; Pred. No. 1.3e-20;
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGACATCATTAATTAAGTGAATTTTCTGGCCCTGAGACTTGACGGAGGCAAGA 60
 DB 442 AGAAGACATCATTAATTAAGTGAATTTTCTGGCCCTGAGACTTGACGGAGGCAAGA 501

QY 61 AGAAGCTCTGGACACCATATGACAG 87
 DB 502 AGAAGCTCTGGACACCATATGACAG 528

RESULT 6
 AC092184 71132 bp DNA linear PRI 12-JUN-2002
 LOCUS Homo sapiens 12 BAC RP11-43817 (Roswell Park Cancer Institute Human
 BAC library) complete sequence.
 ACCESSION AC092184 AC013443
 VERSION AC092184.7 GI:21206067
 KEYWORDS HNG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 71132)
 Muzny, D. M., Adams, C., Adio-Oduola, B., Allousman, F. R., Allen, C.,
 Alsbrooks, S. L., Amaralunga, H. C., Are, J. R., Ayala, M., Banks, T.,
 Barbarella, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,
 Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N. P.,
 Buhay, C., Burch, P., Burkett, C., Burrell, K. L., Byrd, N. C.,
 Caron, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C.,
 Cleveland, C. D., Cox, C., Coyle, M. D., Dahorne, S. R., David, R.,
 Davila, M. L., Davis, C., Davy-Carroll, L., Dederich, D. A.,
 Delaney, K. R., Delgado, O., Denn, A. L., Ding, Y., Dinh, H. H.,
 Douthwaite, K. J., Draper, R., Dugan-Rocha, S., Durbin, K. J.,
 Eacotto, M., Edgar, D., Edwards, C. C., Elhaj, C., Emerling, S.,
 Escotto, M., Falis, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P.,
 Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N.,
 Gill, R., Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S.,
 Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P.,
 Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M.,
 Holloway, C., Hollins, B., Homel, P., Howard, S., Huber, J., Hulik, S.,
 Hume, J., Ioshikhes, I., Jackson, J. E., Jacobson, B., Jia, Y.,
 Johnson, R., Jolivet, S., Joudah, S., Karlsson, S., Kelly, S., Khan, U.,
 King, L., Kovach, J., Kovar, C., Kratochvil, J., Kureishi, A., Landry, N.,
 Leal, B., Lee, E., Lewis, L. C., Lewis, L. J., Li, J., Li, Z., Licharge, O.,
 Lieu, C., Liu, J., Liu, W., Lousseng, H., Lozada, R. J., Lu, X.,
 Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P.,
 Marondel, I., Martin, R., Martindale, A., Martinez, B., Massey, B.,
 Mawhinney, E., McLeod, M. P., Meador, M., Mei, G., Messinger, S.,
 Metzger, M., Miller, A., Miner, G., Miner, Z., Mitchell, T.,

standards - estimated error rate less than 1 per 10,000 bases.
 Reports of lowest quality individual bases and measures of base
 quality are listed below. Description of the metrics can be found
 at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT

Location/Qualifiers

```

1..71132
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-438L7"
/complement(1..1448)
/note="overlaps bases 1..1448 of clone AC092490"
/function="Clone overlap"
187..408
/standard_name="57233"
439..560
/standard_name="92005"
774..881
/rpt_family="MIR"
903..1190
/rpt_family="AluSg"
1191..1213
/rpt_family="AT_rich"
1591..1807
/standard_name="6612"
1744..1819
/standard_name="8198"
1966..2264
/rpt_family="AluSg"
3296..3328
/rpt_family="TTTC)n"
4097..4249
/rpt_family="AluSg"
4360..4654
/rpt_family="AluY"
complement(5140..5262)
/rpt_family="FLAM_C"
5275..5304
/rpt_family="AT_rich"
5305..5563
/rpt_family="AluJo"
5564..5599
/rpt_family="AT_rich"
5644..5733
/rpt_family="MSTD"
5734..6012
/rpt_family="AluSg"
6020..6336
/rpt_family="AluSx"
6337..6629
/rpt_family="MSTD"
6630..6650
/rpt_family="(CAA)n"
complement(6711..6817)
/rpt_family="MIR"
complement(7033..7327)
/rpt_family="AluSx"
7607..7745
/rpt_family="FLAM_A"
7980..8047
/rpt_family="MIR"
complement(8050..8549)
/rpt_family="LTR47A"
8567..8698
/rpt_family="MIR"
complement(8815..9101)
/rpt_family="AluSg"
complement(11227..12175)
/rpt_family="LTR5"

```

misc_feature

STS

STS

repeat_region

repeat_region

repeat_region

STS

STS

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

STS
12625..12900
/standard_name="87433"
repeat_region complement(13547..13664)
/rpt_family="L1MB5"
repeat_region complement(13665..13979)
/rpt_family="L1MB5"
repeat_region complement(13980..14193)
/rpt_family="L1MB5"
repeat_region 14632..14924
/rpt_family="L1MB5"
repeat_region complement(15813..17136)
/rpt_family="L1MB5"
repeat_region complement(17272..17395)
/rpt_family="L1MB5"
repeat_region 17426..17575
/rpt_family="L1MB5"

Query Match 100.0%; Score 87; DB 9; Length 71132;
Best Local Similarity 100.0%; Pred. No. 1.4e-20;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGACCATCATTAATGAAGTGAATTTCTGGCTGAGACTTGACGGAGGCAAGA 60
DB 34990 AGAGACCATCATTAATGAAGTGAATTTCTGGCTGAGACTTGACGGAGGCAAGA 35049
QY 61 AGACACTCTGGACACCACTATGACAG 87
DB 35050 AGACACTCTGGACACCACTATGACAG 35076

RESULT 7
LOCUS BC006296 1828 bp mRNA linear PRI 03-OCT-2003
DEFINITION Homo sapiens activation-induced cytidine deaminase, mRNA (cdna)
ACCESSION BC006296
VERSION BC006296.2 GI:33871601
KEYWORDS MGC.
ORGANISM Homo sapiens (human)
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
1 (bases 1 to 1828)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Dietzenko, L., Marusik, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, K.B., Bonaldo, M.F., Casavant, T.L.,
Scheet, T.E., Brownstein, M.J., Udell, T.B., Toshiyuki, S.,
Abramson, R.D., Wollahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Mizny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalls, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL
MEDLINE
22388257
PUBMED
12477932
REFERENCE
2 (bases 1 to 1828)
Straussberg, R.
Direct Submission
Submitted (09-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:13623400.
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Louis Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,
Dietzen, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Larc, P., Legaspi, R.,
Maduro, Q.L., Masello, C., Maskell, B., Mastrian, S.D., McCloskey, D.C.,
McDowell, J., Pearson, R., Santipol, S., Thomas, P.J., Touchman, J.W.,
Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAL Plate: 17 Row: a Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 10190699.

FEATURES

source

1..1828
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:12911 IMAGE:4054915"
/issue_type="Primary B-Cells from Tonsils"
/clone_lib="NIH MGC 46"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1..1828
/gene="AICDA"
/note="synonyms: AID, HIGM2, CDA2, ARP2"
/db_xref="LocusID:57379"
/db_xref="MIM:605257"
77..673
/codon_start=1
/product="activation-induced cytidine deaminase"
/protein_id="AA06296.1"
/db_xref="GI:13623401"
/db_xref="LocusID:57379"
/translation="MDSILNMRKFLYQKVRNKAQRRETYLCVVRSDSATSFL
DGYLRNKGCHVEFLFRTYISDNDLDRGRCRYRVFMSKSCYCAHVDLFRGN
NLSLIFARLFLFCDRRAEPGLRLRRAAGQILIMPKDFYCMNTFVNHETFK
AMEGHENSVRSLRDLRLILFLYEVDDLDRALFLIG."

gene

CDS

ORIGIN
Query Match 96.6%; Score 84; DB 9; Length 1828;
Best Local Similarity 100.0%; Pred. No. 1.5e-19;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GAACATCATTAATGAAGTGAATTTCTGGCTGAGACTTGACGGAGGCAAGA 63
DB 1 GAACATCATTAATGAAGTGAATTTCTGGCTGAGACTTGACGGAGGCAAGA 60
QY 64 CACTTGACACCACTATGACAG 87
DB 61 CACTTGACACCACTATGACAG 84
RESULT 8
LOCUS AB040431 2791 bp mRNA linear PRI 03-OCT-2000
DEFINITION Homo sapiens AID mRNA for activation-induced cytidine deaminase,
complete CDS.
ACCESSION AB040431
VERSION AB040431.1 GI:9988409

KEYWORDS AID; activation-induced cytidine deaminase; Human AID.
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (sites)
 AUTHORS Muto, T., Muramatsu, M., Taniguchi, M., Kinoshita, K. and Honjo, T.
 TITLE Isolation, tissue distribution, and chromosomal localization of the human activation-induced cytidine deaminase (AID) gene
 JOURNAL Genomics 68 (1), 85-88 (2000)
 MEDLINE 20408950
 PUBMED 10950930
 REFERENCE 2 (sites)
 AUTHORS Revy, P., Muto, T., Levy, Y., Gelsmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Lageat, R., Gennery, A., Tezcan, I., Ersoy, F., Kayserili, H., Ugazio, A. G., Brousse, N., Muramatsu, M., Notarangelo, L. D., Kinoshita, K., Honjo, T., Fischer, A. and Durandy, A.
 TITLE Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2)
 JOURNAL Cell 102 (5), 565-575 (2000)
 MEDLINE 20460541
 PUBMED 11007475
 REFERENCE 3 (bases 1 to 2791)
 AUTHORS Muto, T., Muramatsu, M., Taniguchi, M., Kinoshita, K. and Honjo, T.
 TITLE Direct Submission
 SUBMITTED (18-MAR-2000) Tasuku Honjo, Kyoto University, Department of Medical Chemistry, Faculty of Medicine, Yoshida, Sakyo-ku, Kyoto, Kyoto 606-8501, Japan (E-mail: honjo@imf.med.kyoto-u.ac.jp, Tel:81-75-753-4371(ex.4371), Fax:81-75-753-4388)
 FEATURES
 source Location/Qualifiers
 1..2791
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1..2791
 /gene="AID"
 77..673
 /gene="AID"
 /codon_start=1
 /product="activation-induced cytidine deaminase"
 /protein_id="BAB12721.1"
 /db_xref="GI:9988410"
 /translation="MDSILMRKFLYQFKNRKAKGRREYLCVVRPSATSFSLDFGLKNGCHEVLEFLRYSIDWDLPKRCRTVTPFSWSPCDCAHVADPFRGNPNSLRFTARLYFCEDRKAPRGLRLHAGVQALIMPFKYCYQWTFVNHRTKXAWEGHENSVRSLRQLRILPLVEVDLRDAFRLGL"
 ORIGIN
 Query Match 96.6%; Score 84; DB 9; Length 2791;
 Best Local Similarity 100.0%; Pred. No. 1.5e-19;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 GAACATCATTAATGAGTATTTCTGCGCTGAGACTTGAGGAGGACGAAGA 63
 DB 1 GAACATCATTAATGAGTATTTCTGCGCTGAGACTTGAGGAGGACGAAGA 60
 QY 64 CACTTGACACCACTATGAGACG 87
 DB 61 CACTTGACACCACTATGAGACG 84
 RESULT 9
 AB091291 1767 bp DNA linear ROD 19-SEP-2003
 LOCUS flanking region and partial cde.
 DEFINITION
 ACCESSION AB091291
 VERSION AB091291.1 GI:34447115
 KEYWORDS Mus musculus (house mouse)
 SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1
 AUTHORS Gonda, H., Sugai, M., Nambu, Y., Katakai, T., Agata, Y., Yokota, Y. and Shimizu, A.
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 1767)
 REFERENCE Gonda, H., Sugai, M., Nambu, Y., Katakai, T., Agata, Y., Yokota, Y. and Shimizu, A.
 TITLE Direct Submission
 SUBMITTED (04-SEP-2002) Hiroyuki Gonda, Center for Molecular Biology and Genetics; 53 Shogoin-kawahara-cho, Sakyo-ku, Kyoto 606-8507, Japan (E-mail: hgonda@virius.kyoto-u.ac.jp, Tel:81-75-751-4189, Fax:81-75-751-4190)
 FEATURES
 source Location/Qualifiers
 1..1767
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="6"
 <1..1759
 /note="5' flanking region"
 1760..1767
 /gene="AID"
 1760..1767
 /gene="AID"
 /codon_start=1
 /product="activation-induced cytidine deaminase"
 /protein_id="BAC84974.1"
 /db_xref="GI:34447116"
 /db_xref="LOCUSID:11628"
 /translation="MD"
 ORIGIN
 Query Match 47.1%; Score 41; DB 10; Length 1767;
 Best Local Similarity 75.3%; Pred. No. 0.00075;
 Matches 64; Conservative 0; Mismatches 20; Indels 1; Gaps 1;
 QY 4 GAACATCATTAATGAGTATTTCTGCGCTGAGACTTGAGGAGGACGAAGAAG 62
 DB 1683 GAACATCATTAATGAGTATTTCTGCGCTGAGACTTGAGGAGGACGAAGAAG 1742
 QY 63 ACACTTGACACCACTATGAGACG 87
 DB 1743 TCAAGCTTGACACCACTATGAGACG 1767
 RESULT 10
 BD016828 2440 bp DNA linear PAT 27-AUG-2002
 LOCUS Novel cytidine deaminase.
 DEFINITION
 ACCESSION BD016828
 VERSION BD016828.1 GI:22558004
 KEYWORDS JP 2001245669-A/1.
 SOURCE Mus musculus (house mouse)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2440)
 AUTHORS Honjo, T. and Muramatsu, M.
 TITLE Novel cytidine deaminase
 JOURNAL Patent: JP 2001245669-A 1 11-SEP-2001;
 JAPAN TOBACCO INC, TASUKU HONJO
 COMMENT
 OS Mus musculus (mouse)
 PN JP 2001245669-A/1
 PD 11-SEP-2001
 PF 28-MAR-2000 JP 2000092981
 PI TASUKU HONJO, MASAMICHI MURAMATSU
 PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12, A61P17/00,
 PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC C12N1/21,
 PC C12N5/10, C12N9/78, C12P21/02, C12P21/08, C12N1/21, C12N1/19, PC

FC12N5/10,C12R1:91),C12N5/00,C12N5/00,(C12N5/00,C12R1:91) CC
FH Key Location/Qualifiers
FT CDS (93) . (689)
FT 5'UTR (1) . (192)
FT 3'UTR (680) . (2440).
FEATURES
source
1. .2440
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

Query Match 47.1%; Score 41; DB 6; Length 2440;
Best Local Similarity 75.3%; Pred. No. 0.00076;
Matches 64; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 4 GAACCATCTATTAAATGAGATTCTTGCCCTGAGACTTGACAGGAGCAAG-AGA 62
DB 16 GAAGCAGCCCTTGCTTGAAGCAAGCTTCTTGCCCTTAAGACTTGAGGAGCAAGAAAG 75
QY 63 ACACCTGTGACACCACTATGACAG 87
DB 76 TCACGCTGAGACCACTATGACAG 100

RESULT 11
AF132979 2440 bp mRNA linear ROD 05-AUG-1999
LOCUS AF132979
DEFINITION Mus musculus activation-induced cytidine deaminase (AID) mRNA,
complete cds.
VERSION AF132979
KEYWORDS AF132979.1 GI:5305727
SOURCE Mus musculus
ORGANISM Mus musculus (house mouse)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 2440)
Muraiatsu,M., Sankaranand,V.S., Anant,S., Sugai,M., Kinoshita,K.,
Davidson,N.O. and Honjo,T.
Specific expression of activation-induced cytidine deaminase (AID),
a novel member of the RNA-editing deaminase family in germinal
center B cells
JOURNAL J. Biol. Chem. 274 (26), 18470-18476 (1999)
MEDLINE 99303612
PubMed 10373455
REFERENCE 2 (bases 1 to 2440)
AUTHORS Muraiatsu,M. and Honjo,T.
TITLE Direct Submission
Submitted (04-MAR-1999) Medical Chemistry, Kyoto University, Konoe
Yoshida Sakyo-ku, Kyoto 606-8501, Japan
FEATURES
source
1. .2440
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1. .2440
/gene="AID"
93. .689
/gene="AID"
/note="APOBEC-1 homologue"
/codon_start=1
/product="activation-induced cytidine deaminase"
/protein_id="AAD41793.1"
/db_xref="GI:5305728"
/translation="MDSLMKQKPKLYHKVNRMAKGRHETLYCVVRRPSATSGSL
DGHKLNKSGCHEVLEFLRYSIDWDIDPGRKRVTFWSMSTCYCAHVAEFLKWP
NSLRIFTRALFYCEDRKAPGELRLHAGVQIGIMTFKDYFQWNTFVNRERTFK
AAGELHNSVRLRLRRLRLPLLYEVDLRDAFRLMGF"

Query Match 47.1%; Score 41; DB 10; Length 2440;
Best Local Similarity 75.3%; Pred. No. 0.00076;

Matches 64; Conservative 0; Mismatches 20; Indels 1; Gaps 1;
QY 4 GAACCATCTATTAAATGAGATTCTTGCCCTGAGACTTGACAGGAGCAAG-AGA 62
DB 16 GAAGCAGCCCTTGCTTGAAGCAAGCTTCTTGCCCTTAAGACTTGAGGAGCAAGAAAG 75
QY 63 ACACCTGTGACACCACTATGACAG 87
DB 76 TCACGCTGAGACCACTATGACAG 100

RESULT 12
AC119975 178130 bp DNA linear HTG 14-MAR-2003
LOCUS AC119975
DEFINITION Mus musculus clone RP24-483K3, WORKING DRAFT SEQUENCE, 8 unordered
pieces.
VERSION AC119975 GI:28951252
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 178130)
JOURNAL Unpublished
AUTHORS Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Bouhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,D.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., MacDonald,P., Major,J., Margulis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Menus,L.,
Mhova,T., Mlenga,C.H., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,M., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE
AUTHORS 3 (bases 1 to 178130)
Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Baerlein,V., Bloom,T.,
Boguslavsky,L., Bouhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Halez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Liu,A., Mabbitt,R., Maclean,C.,
MacDonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrum,J., Menus,L., Mhova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,

TITLE
JOURNAL
COMMENT

Wyma, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (14-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 14, 2003 this sequence version replaced gi:28195895.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WTR

Web site: http://www.seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: 125744

Center clone name: 483 K.3

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 176301 bases at least Q40

Consensus quality: 177194 bases at least Q20

Insert size: 176000; agarose-fp

Insert size: 177430; sum-of-contigs

Quality coverage: 9.0 in Q20 bases; agarose-fp

Quality coverage: 8.9 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 11982: contig of 11982 bp in length
* 11983 12082: gap of 100 bp in length
* 12083 13278: contig of 1196 bp in length
* 13279 13378: gap of 100 bp
* 13379 15640: contig of 2262 bp in length
* 15641 15740: gap of 100 bp
* 15741 22800: contig of 7060 bp in length
* 22801 22901: gap of 100 bp
* 22901 101338: contig of 78438 bp in length
* 101339 133612: contig of 32174 bp in length
* 133613 133712: gap of 100 bp
* 133713 173588: contig of 39876 bp in length
* 173589 178130: gap of 100 bp
* 178131 178130: contig of 4442 bp in length.

FEATURES

source

1. 178130

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone="RP24-483K3"

/clone_1ib="RP24-483K3 Male Mouse BAC"

1. 11982

/note="assembly_fragment"

clone_end:SP6

vector_side:left"

12083. 13278

/note="assembly_fragment"

13379. 15640

/note="assembly_fragment"

15741. 22800

/note="assembly_fragment"

22901. 101338

/note="assembly_fragment"

101439. 133612

/note="assembly_fragment"

133713. 173588

ORIGIN

Query Match 47.1%; Score 41; DB 2; Length 178130;
Best Local Similarity 75.3%; Pred. No. 0.00092;
Matches 64; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 4 GAACATATTATTTGAAGTGAATTTTCGCTGACCTTGACGAGGACGAG-ANG 62
DB 138666 GAAGAGCGCTTCTGCTGAAGAGAGCTTCTTGCTTAAGCTTGAAGGAGTCAAGAAAG 138607
QY 63 AACCTGTGACGACGACCTATGACGACG 87
DB 138606 TCACCTGTGACGACGATATGACGACG 138582

RESULT 13
LOCUS AC094826/c
DEFINITION Rattus norvegicus clone CH230-2013, WORKING DRAFT SEQUENCE.
ACCESSION AC094826
VERSION AC094826.4 GI:30466489
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 241757)
Muzny, D., Marie, Metzker, M., Lee, A., Abruzzo, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alibek, S., Amin, A., Angiano, D.,
Anyalebech, V., Ayogaji, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceaar, H., Centes, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
David, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Drapper, H., Dugan-Rocha, S., Dunn, A., Durkin, K., Duval, B., Evans, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C., A., Falls, I., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgiev, G., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunnarsson, P., Healand, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, T., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S., L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulik, S., Hume, J., Ideblid, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Joliver, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowals, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, J., Lopez, J.,
Lorenz, L., Louised, H., Lozada, R., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, K., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milojevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munday, M., Murphy, M., Nat, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwackemele, O., Okunolu, G., Olarnunagson, A., Pal, S., Parks, K.,
Pastermak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C.,
Plopper, F., Polindexter, A., Popovic, D., Prims, E., Pu, L., L.,
Puzo, M., Quito, J., Rachlin, E., Reeves, K., Resler, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J.,
Sanders, M., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartbeyn, A., Sismon, I., Sitter, C. D., Smaj, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,

Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Trejos, Z., Umani, K., Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wiczak, R., Woodson, H., Wortley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE

Unpublished

2 (bases 1 to 241757)

Worley, K.C.

Direct Submission

Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 241757)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:23101539.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GBMC

Center clone name: CH230-2013

----- Summary Statistics

Assembly program: Atlas;

Consensus quality: 227874 bases at least Q40

Consensus quality: 231202 bases at least Q30

Consensus quality: 233430 bases at least Q20

Estimated insert size: 245502; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 241757: contig of 241757 bp in length.

Location/Qualifiers

1. 241757

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-2013"

1. 3654

/note="wgs_contig"

misc_feature

complement(238399..239261)

misc_feature

complement(238399..239261)

misc_feature

complement(238399..239261)

misc_feature

complement(238399..239261)

misc_feature

complement(238399..239261)

misc_feature

complement(238399..239261)

/note="clone_boundary
clone end:17
site:ECORI
end_sequence:BH289756"

Query Match 39.8%; Score 34.6; DB 2; Length 241757;
Best Local Similarity 70.6%; Pred. No. 0.2;
Matches 60; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

4 GAACATCATTTAATGAGATTTTCTGCGCTGACATCTTGCAGGAGGACAG-AG 62

118777 GAAGCAGCCCTGCTTGAAGCAAGCTCTTGAACCTTGAAGCTTTGAGGAGAAAG 118718

63 ACACCTGAGACACCAATGAGAG 87

118717 TCACGCTGAGAGACCAATGAGAG 118693

RESULT 14
AC109119/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-98A4. *** SEQUENCING IN PROGRESS ***.
AC109119
AC109119.5 GI:22856767
VERSION
HMG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Rattus norvegicus

1 (bases 1 to 259506)
Muzny, D., Wente, M., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amth, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyly, M., Cree, A., D'Souza, L., Davila, M., L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, M., Gunatregoria, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, D., Hernandez, R., Hines, S., Hladun, S. L., Hodgeson, A., Hogues, M., Hollins, B., Howells, S., Hui, Y., Hume, J., Idlebird, D., Jolyet, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuwa, L., Louised, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Mirza, E., Mollenhauer, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundaasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norriss, S., Nwankweme, O., Okunolu, G., Olaprunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, P., Poindexter, A., Popovic, D., Primm, B., Pu, L., L., Puzo, N., Quirroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reich, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shastman, S., Shen, H., Shetty, V., Shvartbeyn, A., Sison, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrell, R., Sosa, O., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,

Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wleczek, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE
UnpublishedJOURNAL
2 (bases 1 to 345098)REFERENCE
Worley, K. C.AUTHORS
TITLE

JOURNAL

Submitted (09-May-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

(bases 1 to 345098)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (15-Nov-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 15, 2002 this sequence version replaced gi:23265477.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GWTN

Center clone name: CH230-3215

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 191191 bases at least Q40

Consensus quality: 194508 bases at least Q30

Estimated insert size: 227256; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 330403: contig of 330403 bp in length
* 330404 330503: gap of unknown length
* 330504 331503: contig of 1100 bp in length
* 331504 331703: gap of unknown length
* 331704 331738: contig of 6035 bp in length
* 331739 337838: gap of unknown length
* 337839 345098: contig of 7260 bp in length.

FEATURES

source

1.345098
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-3215"

misc_feature complement(177532..178431)

/note="clone boundary

clone_end:Sp6

site: end_sequence:BH328475"

179290..181054

/note="wgs_end_extension

clone_end:Sp6"

ORIGIN

Query Match 39.8%; Score 34.6; DB 2; Length 345098;

Best Local Similarity 70.6%; Pred. No. 0.21; Mismatches 24; Indels 1; Gaps 1;

Matches 60; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 4 GAACCATCAATTAATGAAGTAGAGATTCTTGCGCTGAGACTTGAGGAGGCAAG-AAG 62

DB 149860 GAAGCAGCCTTGCTTGAGCAAGCTCCCTTGAACCTTAAGACTTTGAGGAAACAAAGAAAG 149801

QY 63 ACACTTGACACCACTATGACAG 87

DB 149800 TCAAGCTGAGACCGAAATGACAG 149776

Search completed: March 12, 2004, 21:22:29
Job time : 277.899 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 12:26:39 ; Search time 34.8377 Seconds

(without alignments)
10609.014 Million cell updates/sec

Title: US-09-966-880a-11

Perfect score: 87
Sequence: 1 agagacacatcatatgta.....ctggacacacactatgacag 87

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04:*

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 87 | 100.0 | 87 | 3 | AAC55315 Human act |
| 2 | 87 | 100.0 | 2818 | 3 | AAC55312 Human act |
| 3 | 87 | 100.0 | 5514 | 3 | AAC55313 Human act |
| 4 | 87 | 100.0 | 11204 | 3 | AAC55339 Human act |
| 5 | 87 | 100.0 | 11204 | 6 | AB573286 DNA encod |
| 6 | 84 | 96.6 | 1543 | 7 | ABX05468 Human nov |
| 7 | 84 | 96.6 | 2791 | 6 | AB573287 DNA encod |
| 8 | 84 | 96.6 | 2791 | 6 | AB573288 DNA encod |
| 9 | 41 | 47.1 | 2440 | 3 | AAC55307 Mouse act |
| 10 | 30.6 | 35.2 | 42299 | 4 | AAK68932 Human imm |
| 11 | 30 | 34.5 | 328 | 6 | ABQ97746 Mouse ES |
| 12 | 28.8 | 33.1 | 762 | 2 | AAZ16628 Human gen |
| 13 | 28.2 | 32.4 | 958 | 2 | AAV33481 Human dis |
| 14 | 27.8 | 32.0 | 2716 | 9 | ADD48661 Rat gene |
| 15 | 27.6 | 31.7 | 444 | 9 | ADD32481 Human mit |
| 16 | 27.6 | 31.7 | 496 | 4 | AA185721 Human kin |
| 17 | 27.6 | 31.7 | 515 | 8 | AA557351 Human kin |
| 18 | 27.6 | 31.7 | 843 | 2 | AA731065 Human kdp |
| 19 | 27.6 | 31.7 | 1371 | 10 | AD28378 Human sec |
| 20 | 27.4 | 31.5 | 947 | 7 | AA85023 Human sec |
| 21 | 27.4 | 31.5 | 947 | 7 | ACD18949 Novel hum |
| 22 | 27.2 | 31.3 | 5810 | 5 | AA584731 DNA encod |
| 23 | 27.2 | 31.3 | 10497 | 4 | AAK67302 Human imm |

| | | | | | | |
|---|----|------|------|-------|---|--------------------|
| C | 24 | 27.2 | 31.3 | 10497 | 4 | AAK83166 Human imm |
| | 25 | 27 | 31.0 | 389 | 4 | AA186461 Human pol |
| | 26 | 26.4 | 30.3 | 360 | 2 | AAQ61230 Human bra |
| | 27 | 26.2 | 30.1 | 9353 | 7 | AB268800 Nucleotid |
| | 28 | 26.2 | 30.1 | 10597 | 7 | AB268799 Nucleotid |
| | 29 | 26.2 | 30.1 | 10953 | 7 | AB268805 Nucleotid |
| C | 30 | 26.2 | 30.1 | 32154 | 4 | AAJ37068 Human mus |
| | 31 | 26.2 | 30.1 | 32154 | 4 | ABK60056 CDNA enco |
| | 32 | 26 | 29.9 | 26 | 3 | AAC55335 Human act |
| | 33 | 26 | 29.9 | 412 | 6 | AB235592 Human gen |
| | 34 | 26 | 29.9 | 417 | 7 | AB271602 Breast sp |
| | 35 | 26 | 29.9 | 444 | 7 | ACF56780 Rice endo |
| | 36 | 26 | 29.9 | 1328 | 7 | AB271603 Breast sp |
| | 37 | 26 | 29.9 | 2004 | 4 | ABL10849 Drosophi1 |
| C | 38 | 26 | 29.9 | 2382 | 4 | AAH18091 Human cdn |
| | 39 | 26 | 29.9 | 2386 | 9 | ADD29848 Human tum |
| C | 40 | 26 | 29.9 | 2772 | 4 | AAK5945 Human imm |
| | 41 | 25.8 | 29.7 | 543 | 3 | AAK8238 Human sec |
| C | 42 | 25.8 | 29.7 | 543 | 6 | ABK36224 CDNA sequ |
| | 43 | 25.8 | 29.7 | 2506 | 4 | AB12873 Drosophi1 |
| | 44 | 25.8 | 29.7 | 4117 | 4 | AB127156 Drosophi1 |
| | 45 | 25.8 | 29.7 | 4699 | 4 | AB112872 Drosophi1 |

ALIGNMENTS

| | | |
|----------|--|-----------------------|
| RESULT 1 | AAK55315 | standard; DNA; 87 BP. |
| ID | AAC55315 | |
| AC | AAC55315; | |
| XX | | |
| DT | 05-FEB-2001 | (first entry) |
| XX | | |
| DE | Human activation-induced cytidine deaminase exon 1 SEQ ID NO:11. | |
| XX | | |
| KW | Activation-induced cytidine deaminase; AID; cytidine deaminase; | AAK83166 Human imm |
| KW | immune related disease; allergy; allergic disease; antiallergic; | AA186461 Human pol |
| KW | antianaemic; antiaethmatic; ophthalmological; anti-HIV; dermatological; | AAQ61230 Human bra |
| KW | gene therapy; B cell associated immune system disorder; food allergy; | AB268800 Nucleotid |
| KW | immunodeficiency disease; immunoglobulin A deficiency disease; asthma; | AB268799 Nucleotid |
| KW | IGA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis; | AB268805 Nucleotid |
| KW | drug allergy; allergic rhinitis; Rosen disease; Digeorge disease; AIDS; | AAJ37068 Human mus |
| KW | ataxia telangiectasia; common variable immunodeficiency disorder; | ABK60056 CDNA enco |
| KW | major histocompatibility class II deficiency disease; | AAC55335 Human act |
| KW | auto immunodeficiency syndrome; Igg subclass selection disorder; ds. | AB235592 Human gen |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| PN | W0200058480-A1. | |
| XX | | |
| PD | 05-OCT-2000. | |
| XX | | |
| PF | 28-MAR-2000; 2000MO-JP001918. | |
| XX | | |
| PR | 29-MAR-1999; 99JP-00087192. | |
| PR | 24-JUN-1999; 99JP-00178999. | |
| XX | 27-DEC-1999; 99JP-00371382. | |
| XX | | |
| PA | (NIBS) JAPAN TOBACCO INC. | |
| PA | (HONT/) HONTO T. | |
| PI | Honjo T, Muramatsu M; | |
| XX | WPI; 2000-611715/58. | |
| XX | | |
| PT | Nucleic acid encoding activation induced cytidine deaminase, useful as a | |
| PT | target for drug development for immune-related diseases including | |
| XX | allergies. | |
| XX | | |
| PS | Claim 18; Page 150; 174pp; Japanese. | |
| XX | | |

CC The present invention describes an activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has anti-allergic, antianaemic,
 CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IgA) deficiency disease, IgA nephritis, gamma-
 CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, DisGeorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, MHC (major
 CC histocompatibility class II deficiency disease, AIDS (auto
 CC immunodeficiency syndrome), elevated Ige disorder, and Igg subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders. The present sequence represents the
 CC exon 1 genomic DNA sequence of human AID

Query Match 100.0%; Score 87; DB 3; Length 87;
 Best Local Similarity 100.0%; Pred. No. 3.5e-22;

Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAACCATCATTAATTGAGAGATTTCCTGCGCTGAGACTTGACGGAGGCAAGA 60
 1 AGAGAACCATCATTAATTGAGAGATTTCCTGCGCTGAGACTTGACGGAGGCAAGA 60

QY 61 AGACACTCTGGACACCACTATGACAG 87
 61 AGACACTCTGGACACCACTATGACAG 87

RESULT 2

ID AAC55312 standard; cDNA, 2818 BP.

AC AAC55312;

DT 05-FEB-2001 (first entry)

DE Human activation-induced cytidine deaminase encoding cDNA SEQ ID NO:7.

XX Activation-induced cytidine deaminase; AID; cytidine deaminase;
 KM immune related disease; allergy; allergic disease; anti-allergic;
 KM antianaemic; antiasthmatic; ophthalmological; anti-HIV; dermatological;
 KM gene therapy; B cell associated immune system disorder; food allergy;
 KM immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
 KM IgA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;
 KM drug allergy; allergic rhinitis; Rosen disease; DisGeorge disease; AIDS;
 KM ataxia telangiectasia; common variable immunodeficiency disorder;
 KM major histocompatibility class II deficiency disease;
 KM auto immunodeficiency syndrome; Igg subclass selection disorder; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 80..676

FT /tag= a /product= "activation-induced cytidine deaminase"

XX MO200058480-A1.

XX 05-OCT-2000.

XX 28-MAR-2000; 2000MO-JP001918.

XX 29-MAR-1999; 99JP-00087192.

XX 24-JUN-1999; 99JP-00178999.

XX 27-DEC-1999; 99JP-00371382.

XX (NIBS) JAPAN TOBACCO INC.

PA (HONJ/) HONJO T.

XX Honjo T. Muramatsu M.
 PI WPI, 2000-611715/58.
 DR P-PDB; AAB24198.
 XX Nucleic acid encoding activation induced cytidine deaminase, useful as a
 PT target for drug development for immune-related diseases including
 PT allergies.

PS Claim 3; Page 135-139; 174pp; Japanese.

XX The present sequence encodes human activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has anti-allergic, antianaemic,
 CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IgA) deficiency disease, IgA nephritis, gamma-
 CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, DisGeorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, MHC (major
 CC histocompatibility class II deficiency disease, AIDS (auto
 CC immunodeficiency syndrome), elevated Ige disorder, and Igg subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders

QY Sequence 2818 BP; 868 A; 548 C; 626 G; 776 T; 0 U; 0 Other;

Query Match 100.0%; Score 87; DB 3; Length 2818;
 Best Local Similarity 100.0%; Pred. No. 1.1e-21;
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAACCATCATTAATTGAGAGATTTCCTGCGCTGAGACTTGACGGAGGCAAGA 60
 1 AGAGAACCATCATTAATTGAGAGATTTCCTGCGCTGAGACTTGACGGAGGCAAGA 60

QY 61 AGACACTCTGGACACCACTATGACAG 87
 61 AGACACTCTGGACACCACTATGACAG 87

RESULT 3

ID AAC55313 standard; DNA, 5514 BP.

AC AAC55313;

DT 05-FEB-2001 (first entry)

DE Human activation-induced cytidine deaminase genomic DNA SEQ ID NO:9.

XX Activation-induced cytidine deaminase; AID; cytidine deaminase;
 KM immune related disease; allergy; allergic disease; anti-allergic;
 KM antianaemic; antiasthmatic; ophthalmological; anti-HIV; dermatological;
 KM gene therapy; B cell associated immune system disorder; food allergy;
 KM immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
 KM IgA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;
 KM drug allergy; allergic rhinitis; Rosen disease; DisGeorge disease; AIDS;
 KM ataxia telangiectasia; common variable immunodeficiency disorder;
 KM major histocompatibility class II deficiency disease;
 KM auto immunodeficiency syndrome; Igg subclass selection disorder; ds.

XX Homo sapiens.

XX MO200058480-A1.

XX 05-OCT-2000.

XX 28-MAR-2000; 2000MO-JP001918.

PR 29-MAR-1999; 99JP-00087192.
 PR 24-JUN-1999; 99JP-00178999.
 PR 27-DEC-1999; 99JP-00371382.

XX (NIBS) JAPAN TOBACCO INC.
 PA (HONJ) HONJO T.

PI Honjo T, Muramatsu M;

XX WPI; 2000-611715/58.

DR Nucleic acid encoding activation induced cytidine deaminase, useful as a
 PT target for drug development for immune-related diseases including
 PT allergies.

XX Claim 17; Page 142-145; 174pp; Japanese.

XX The present invention describes an activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic,
 CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
 CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, Digorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, MHC (major
 CC histocompatibility class II deficiency disease, AIDS (auto
 CC immunodeficiency syndrome), elevated IGE disorder, and IGG subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders. The present sequence represents a
 CC genomic DNA sequence of human AID

XX SO Sequence 5514 BP; 1709 A; 1045 C; 1134 G; 1623 T; 0 U; 3 Other;

Query Match 100.0%; Score 87; DB 3; Length 5514;

Best Local Similarity 100.0%; Pred. No. 1.4e-21;
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAACCATCATTTAATTGAGTGAATTTTCTGGCCTGAGACTTGACGGAGGCAAGA 60

DB 1032 AGAGAACCATCATTTAATTGAGTGAATTTTCTGGCCTGAGACTTGACGGAGGCAAGA 1091

QY 61 AGACACTCTGGACACCACTATGACAG 87

DB 1092 AGACACTCTGGACACCACTATGACAG 1118

RESULT 4

AAC55339 standard; DNA; 11204 BP.

XX AAC55339;

DT 05-FEB-2001 (first entry)

XX Human activation-induced cytidine deaminase genomic DNA SEQ ID NO:35.

XX Activation-induced cytidine deaminase; AID; cytidine deaminase;
 KW immune related disease; allergy; allergic disease; antiallergic;
 KW antianemic; antiasthmatic; ophthalmological; anti-HIV; dermatological;
 KW gene therapy; B cell associated immune system disorder; food allergy;
 KW immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
 KW IGA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;
 KW drug allergy; allergic rhinitis; Rosen disease; Digorge disease; AIDS;
 KW ataxia telangiectasia; common variable immunodeficiency disorder;
 KW major histocompatibility class II deficiency disease;
 KW auto immunodeficiency syndrome; IGG subclass selection disorder; ds.
 XX Homo sapiens.

PN WO200058480-A1.

XX 05-OCT-2000.

XX 28-MAR-2000; 2000MO-JP001918.

XX 29-MAR-1999; 99JP-00087192.

PR 24-JUN-1999; 99JP-00178999.

PR 27-DEC-1999; 99JP-00371382.

XX (NIBS) JAPAN TOBACCO INC.

PA (HONJ) HONJO T.

PI Honjo T, Muramatsu M;

XX WPI; 2000-611715/58.

DR Nucleic acid encoding activation induced cytidine deaminase, useful as a

PT target for drug development for immune-related diseases including

PT allergies.

XX Claim 17; Page 163-170; 174pp; Japanese.

XX The present invention describes an activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic,
 CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
 CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, Digorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, MHC (major
 CC histocompatibility class II deficiency disease, AIDS (auto
 CC immunodeficiency syndrome), elevated IGE disorder, and IGG subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders. The present sequence represents a
 CC genomic DNA sequence of human AID

XX SO Sequence 11204 BP; 3305 A; 2273 C; 2373 G; 3253 T; 0 U; 0 Other;

Query Match 100.0%; Score 87; DB 3; Length 11204;

Best Local Similarity 100.0%; Pred. No. 1.8e-21;
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAACCATCATTTAATTGAGTGAATTTTCTGGCCTGAGACTTGACGGAGGCAAGA 60

DB 442 AGAGAACCATCATTTAATTGAGTGAATTTTCTGGCCTGAGACTTGACGGAGGCAAGA 501

QY 61 AGACACTCTGGACACCACTATGACAG 87

DB 502 AGACACTCTGGACACCACTATGACAG 528

RESULT 5

AB573286 standard; DNA; 11204 BP.

XX AB573286;

DT 04-DEC-2002 (first entry)

XX DNA encoding human translocation del(12p) protein #1.

XX Chromosome aberration; oncogenic fusion protein; cancer;
 KW proliferative disease; cellular protein isoform; heat shock protein 90;
 KW HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
 KW T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
 KW acute myeloid leukaemia; ALL; chronic myelomonocytic leukaemia; CMML;
 KW acute lymphoblastic leukaemia; ALL; NHL; solid tumour;
 KW papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;

KM rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.
 XX Homo sapiens.
 XX WO200269900-A2.
 XX
 XX 12-SEP-2002.
 XX
 XX 01-MAR-2002; 2002WO-US006518.
 XX
 XX 01-MAR-2001; 2001US-0272751P.
 XX
 XX (CONF-) CONFORMA THERAPEUTICS CORP.
 XX
 XX Fritz LC, Burrows FJ;
 XX WPI; 2002-698710/75.
 XX P-PSDB; ABG95082.
 XX
 XX Treating genetically-defined disease associated with chromosomal
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
 PT diseases, involves administering an inhibitor of heat shock protein 90.
 XX
 XX Disclosure; Page 242-245; 389pp; English.
 XX
 XX The invention describes a method of treating genetically-defined disease
 CC associated with chromosomal aberrations yielding oncogenic fusion
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous
 CC cell population, treating proliferative diseases associated with mutant
 CC protein or cellular protein isoforms (II) dependent on heat shock protein
 CC (HSP)-90, or selectively treating cells expressing (II) involving
 CC administering HSP90-inhibitor. The method is useful for treating
 CC genetically-defined disease with chromosomal aberration yielding
 CC oncogenic fusion protein, treating cancerous cells containing fusion
 CC protein in heterogeneous cell population, treating proliferative disease
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
 CC p53), or selectively treating cells expressing mutant protein or cellular
 CC protein isoform in a patient heterozygous for (II). The method is useful
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CML,
 CC or a disease characterised by a solid tumour such as papillary thyroid
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
 CC synovial sarcoma. The method is also useful for treating viral
 CC infections. This represents the DNA sequence of a chromosome aberration
 SQ Sequence 11204 BP; 3305 A; 2273 C; 2373 G; 3253 T; 0 U; 0 Other;
 Query Match 100.0%; Score 87; DB 6; Length 11204;
 Best Local Similarity 100.0%; Pred. No. 1.8e-21;
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAGAGACCATCATTAATGAGATTTCTGCGCTGAGACTTGAGGAGGAGGAGA 60
 DB 442 AAGAGACCATCATTAATGAGATTTCTGCGCTGAGACTTGAGGAGGAGGAGA 501
 QY 61 AGACACTCTGAGACCACTATGACAG 87
 DB 502 AGACACTCTGAGACCACTATGACAG 528
 RESULT 6
 ABX05468
 ID ABX05468 standard; cDNA; 1543 BP.
 XX
 XX ABX05468;
 XX
 XX 17-JAN-2003 (first entry)
 XX
 XX Human novel polynucleotide #483.
 XX
 XX Human; gene; ss; genetic disorder; gene mapping; medical imaging; cancer;
 KM neurodegenerative disorder; lymphoid cell disorder; osteoporosis;

KM Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
 KM osteoarthritis; periodontal disease; liver fibrosis; viral infection;
 KM fungal infection; bacterial infection; autoimmune disease; diabetes;
 KM atopic dermatitis.
 XX Homo sapiens.
 XX WO200274961-A1.
 XX
 XX 26-SEP-2002.
 XX
 XX 14-MAR-2002; 2002WO-US005109.
 XX
 XX 15-MAR-2001; 2001US-00810173.
 XX
 XX (HYSEQ) HYSEQ INC.
 XX
 XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QH, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehtman T, Wang J, Wang D, Drmanac RT;
 XX WPI; 2003-040556/03.
 DR P-PSDB; ABU00390.
 XX
 XX New isolated polypeptides and polynucleotides, useful for preventing,
 PT treating or ameliorating medical conditions, such as cancer,
 PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
 PT disorders, and infections.
 XX
 XX Claim 1; SEQ ID NO 483; 235pp; English.
 XX
 XX The invention relates to human polynucleotides and the polypeptides they
 CC encode. The polynucleotides and polypeptides are useful in diagnostics,
 CC forensics, gene mapping, medical imaging, identification of mutations
 CC responsible for genetic disorders or other traits, assessing biodiversity
 CC and producing many other types of data and products dependent on DNA and
 CC amino acid sequences. They are also useful for preventing, treating or
 CC ameliorating medical conditions, such as cancer, neurodegenerative
 CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
 CC disorders (e.g. osteoporosis, osteoarthritis, bone degenerative disorders,
 CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
 CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
 CC Sequences ABX04986-ABX05511 represent human polynucleotides of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification but is based on sequence information supplied
 CC by the European Patent Office
 XX
 XX Sequence 1543 BP; 428 A; 334 C; 352 G; 429 T; 0 U; 0 Other;
 SQ
 Query Match 96.6%; Score 84; DB 7; Length 1543;
 Best Local Similarity 100.0%; Pred. No. 1.2e-20;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 GAACCATCATTAATGAGATTTCTGCGCTGAGACTTGAGGAGGAGGAGA 63
 DB 2 GAACCATCATTAATGAGATTTCTGCGCTGAGACTTGAGGAGGAGGAGA 61
 QY 64 CACTCTGAGACCACTATGACAG 87
 DB 62 CACTCTGAGACCACTATGACAG 85
 RESULT 7
 ABS73287
 ID ABS73287 standard; DNA; 2791 BP.
 XX
 XX ABS73287;
 XX
 XX 04-DEC-2002 (first entry)
 XX
 XX DNA encoding human translocation del(12p) protein #2.
 XX
 XX Chromosome aberration; oncogenic fusion protein; cancer;

KM proliferative disease; cellular protein isoform; heat shock protein 90;
 KM HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
 KM T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
 KM acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
 KM acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
 KM papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
 KM rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.
 XX
 OS Homo sapiens.
 PN WO200269900-A2.
 PD 12-SEP-2002.
 PF 01-MAR-2002; 2002WO-US06518.
 PR 01-MAR-2001; 2001US-0272751P.
 PA (CONF-) CONFORMA THERAPEUTICS CORP.
 PI Fritz LC, Burrows FU;
 PI WPI; 2002-698710/75.
 DR P-PSDB; ABG95083.
 PT Treating genetically-defined disease associated with chromosomal
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
 PT diseases, involves administering an inhibitor of heat shock protein 90.
 XX
 PS Disclosure; Page 246-247; 389pp; English.
 XX
 CC The invention describes a method of treating genetically-defined disease
 CC associated with chromosomal aberrations yielding oncogenic fusion
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous
 CC cell population, treating proliferative diseases associated with mutant
 CC protein or cellular protein isoforms (II) dependent on heat shock protein
 CC (HSP)-90, or selectively treating cells expressing (II) involving
 CC administering HSP90-inhibitor. The method is useful for treating
 CC genetically-defined disease with chromosomal aberration yielding
 CC oncogenic fusion protein, treating cancerous cells containing fusion
 CC protein in heterogeneous cell population, treating proliferative disease
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
 CC p53), or selectively treating cells expressing mutant protein or cellular
 CC protein isoform in a patient heterozygous for (II). The method is useful
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, AML, NHL and CMML,
 CC or a disease characterised by a solid tumour such as papillary thyroid
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
 CC synovial sarcoma. The method is also useful for treating viral
 CC infections. This represents the DNA sequence of a chromosome aberration
 XX
 SQ Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;
 Query Match 96.6%; Score 84; DB 6; Length 2791;
 Best Local Similarity 100.0%; Pred. No. 1.4e-20;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 GAACCATTAATTGAAGTGAATTTTCTGGCCCTGAGACTTGACGGAGGAAGA 63
 DB 1 GAACCATTAATTGAAGTGAATTTTCTGGCCCTGAGACTTGACGGAGGAAGA 60
 QY 64 CACTTGACACCACTATGACAG 87
 DB 61 CACTTGACACCACTATGACAG 84

RESULT 8
 AB873288
 ID AB873288 standard; DNA; 2791 BP.
 XX
 AC AB873288;
 XX

DT 04-DEC-2002 (first entry)
 XX
 XX DNA encoding human translocation del(12p) protein #3.
 DE
 XX Chromosome aberration; oncogenic fusion protein; cancer;
 KM proliferative disease; cellular protein isoform; heat shock protein 90;
 KM HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
 KM T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
 KM acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
 KM acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
 KM papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
 KM rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.
 XX
 OS Homo sapiens.
 PN WO200269900-A2.
 PD 12-SEP-2002.
 PF 01-MAR-2002; 2002WO-US06518.
 PR 01-MAR-2001; 2001US-0272751P.
 PA (CONF-) CONFORMA THERAPEUTICS CORP.
 PI Fritz LC, Burrows FU;
 PI WPI; 2002-698710/75.
 DR P-PSDB; ABG95084.
 PT Treating genetically-defined disease associated with chromosomal
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
 PT diseases, involves administering an inhibitor of heat shock protein 90.
 XX
 PS Disclosure; Page 248-249; 389pp; English.
 XX
 CC The invention describes a method of treating genetically-defined disease
 CC associated with chromosomal aberrations yielding oncogenic fusion
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous
 CC cell population, treating proliferative diseases associated with mutant
 CC protein or cellular protein isoforms (II) dependent on heat shock protein
 CC (HSP)-90, or selectively treating cells expressing (II) involving
 CC administering HSP90-inhibitor. The method is useful for treating
 CC genetically-defined disease with chromosomal aberration yielding
 CC oncogenic fusion protein, treating cancerous cells containing fusion
 CC protein in heterogeneous cell population, treating proliferative disease
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
 CC p53), or selectively treating cells expressing mutant protein or cellular
 CC protein isoform in a patient heterozygous for (II). The method is useful
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, AML, NHL and CMML,
 CC or a disease characterised by a solid tumour such as papillary thyroid
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
 CC synovial sarcoma. The method is also useful for treating viral
 CC infections. This represents the DNA sequence of a chromosome aberration
 XX
 SQ Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;
 Query Match 96.6%; Score 84; DB 6; Length 2791;
 Best Local Similarity 100.0%; Pred. No. 1.4e-20;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 GAACCATTAATTGAAGTGAATTTTCTGGCCCTGAGACTTGACGGAGGAAGA 63
 DB 1 GAACCATTAATTGAAGTGAATTTTCTGGCCCTGAGACTTGACGGAGGAAGA 60
 QY 64 CACTTGACACCACTATGACAG 87
 DB 61 CACTTGACACCACTATGACAG 84

RESULT 9

AAc5307
ID AAC5307 standard; cDNA; 2440 BP.
XX
AC AAC5307;
XX
DT 05-FEB-2001 (first entry)
XX
DE Mouse activation-induced cytidine deaminase encoding cDNA SEQ ID NO:1.
XX
XX Activation-induced cytidine deaminase; AID; cytidine deaminase;
KW immune related disease; allergy; allergic disease; antiallergic;
KW antineoplastic; ophthalmological; anti-HIV; dermatological;
KW gene therapy; B cell associated immune system disorder; food allergy;
KW immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
KW IGA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis;
KW drug allergy; allergic rhinitis; Rosen disease; DiGeorge disease; AIDS;
KW ataxia telangiectasia; common variable immunodeficiency disorder;
KW major histocompatibility class II deficiency disease;
KW auto immunodeficiency syndrome; IgG subclass selection disorder; ss.
XX
XX Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 93..689
FT CDS /*tag= a
FT CDS /product= "activation-induced cytidine deaminase"
XX
XX NC020058460-A1.
XX
PD 05-OCT-2000.
XX
PF 28-MAR-2000; 2000MO-JP001918.
XX
XX 29-MAR-1999; 99JP-00087192.
XX 24-JUN-1999; 99JP-00178999.
XX 27-DEC-1999; 99JP-00371382.
XX
PA (NIBS) JAPAN TOBACCO INC.
XX (HONU/) HONTO T.
XX
XX Honjo T, Muramatsu M;
XX
XX WPI; 2000-611715/58.
XX P-PSDB; AAB24197.
XX
XX Nucleic acid encoding activation induced cytidine deaminase, useful as a
XX target for drug development for immune-related diseases including
XX allergies.
XX
XX PS Claim 3; Page 126-130; 174pp; Japanese.
XX
XX The present sequence encodes mouse activation-induced cytidine deaminase
XX (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
XX cytidine activity similar to APOBEC-1. AID has antiallergic, antineoplastic,
XX antineoplastic, ophthalmological, anti-HIV and dermatological activities,
XX and can be used in gene therapy. AID polynucleotides are useful in
XX methods for identifying drugs for the treatment of B cell associated
XX immune system disorders, immunodeficiency diseases and allergies. Such as
XX immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
XX globulinaemia, atopic dermatitis, allergic colitis, asthma, food allergy,
XX drug allergy, allergic rhinitis, Rosen disease, DiGeorge disease, ataxia
XX telangiectasia, common variable immunodeficiency disorder, MHC (major
XX histocompatibility class II deficiency disease, AIDS (auto
XX immunodeficiency syndrome), elevated IgG disorder, and IgG subclass
XX selection disorder. The DNA sequences encoding AID may be used for gene
XX therapy and the antibodies to the AID protein may be used for diagnosis
XX and treatment of these disorders
XX
XX Sequence 2440 BP; 706 A; 546 C; 551 G; 636 T; 0 U; 1 Other;
XX
XX Query Match 47.1%; Score 41; DB 3; Length 2440;
XX Best Local Similarity 75.3%; Pred. No. 0.00011;
XX Matches 64; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

CY 4 GAACCATCATTAATTTGAAGTGAATTTTCTGGCCCTGAGACTTTCAGGAGGCAAG-ANG 62
DB 16 GAAGCAGCCCTGCTTGAAGCAGACTTCCTTGGCTTAAGACTTGGAGAGTCAAGAAAG 75
CY 63 ACACTGAGACACACCATATGAGACAG 87
DB 76 TCAAGCTGAGACCGATATGAGACAG 100
RESULT 10
AAK68932
ID AAK68932 standard; DNA; 42299 BP.
XX
XX AAK68932;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23744.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX NC020057182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001MO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205151P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226868P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.

XX Mouse ES cell related CDNA SEQ ID NO 1014.
 XX AA216628/c
 XX ID
 XX AA216628 standard; CDNA; 762 BP.
 XX AC
 XX AA216628;
 XX DT
 XX 12-OCT-1999 (first entry)
 XX DE
 XX Human gene expression product CDNA sequence SEQ ID NO:4098.
 XX KW
 XX Human; gene; gene expression product; diagnosis; therapy; probe;
 XX KW
 XX detection; mapping; tissue typing; profiling; forensic; cancer;
 XX KW
 XX genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
 XX OS
 XX Homo sapiens.
 XX PN
 XX WO9338972-A2.
 XX PD
 XX 05-AUG-1999.
 XX PF
 XX 28-JAN-1999; 99WC-US001619.
 XX PR
 XX 28-JAN-1999; 98US-0072810P.
 XX PR
 XX 24-FEB-1998; 98US-007354P.
 XX PR
 XX 31-MAR-1998; 98US-0080114P.
 XX PR
 XX 03-APR-1998; 98US-0080515P.
 XX PR
 XX 03-APR-1998; 98US-0080666P.
 XX PR
 XX 21-OCT-1998; 98US-0105234P.
 XX PR
 XX 28-OCT-1998; 98US-0105677P.
 XX PA
 XX (CHIR) CHIRON CORP.
 XX PA
 XX (HYSE-) HYSEQ INC.
 XX PI
 XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 XX PI
 XX Reinhard C, Glese K, Randazzo F, Kennedy GC, Pot D, Kaseam A;
 XX PI
 XX Lamson G, Drmanac R, Ckvenjakov R, Dickson M, Drmanac S, Labat I;
 XX PI
 XX Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Grain B;
 XX XX
 XX WPI; 1999-494092/41.
 XX DR
 XX Novel human genes and their expression products which are differentially
 XX PT
 XX expressed in different cell types.
 XX PS
 XX Claim 1; Page 1942; 2479pp; English.

XX The invention relates to isolated murine polynucleotides (I) comprising a
 XX contiguous stretch of at least about 60 nucleotides of a sequence
 XX (AB096733-AB098191) chosen from 1461 OMNIBANK gene trapped sequences
 XX (GTSs). The novel genes can be used in a process to identify novel
 XX polynucleotide sequences by comparing them to the novel gene sequences.
 XX The novel genes and cells are useful in functional genomic analysis and
 XX in the discovery and development of new therapeutic and diagnostic agents
 XX and methods. (I) is useful for identifying the coding regions of the
 XX murine genome, to isolate cDNAs, genomic clones or full-length
 XX genes/polynucleotides or homologues, heterologues, paralogues or
 XX orthologues that are capable of hybridizing to one or more of the GTSs
 XX under stringent conditions. (I) can be incorporated into a phage display
 XX system that can be used to screen for proteins or other ligands that are
 XX capable of binding an amino acid sequence encoded by an oligonucleotide
 XX or polynucleotide sequence in at least one of the GTS sequences. (I) is
 XX useful in arrays, such as gene chips, to identify and characterize
 XX temporal and tissue specific gene expression, to identify the gene of
 XX interest from many sources and for genetic manipulations such as
 XX antisense inhibition and gene targeting. Decreasing the level of
 XX expression of (I) and/or down regulating the activity of peptides or
 XX proteins encoded by (I) is useful for treating development and cell
 XX differentiation disorders. Note: The sequence data for this patent did
 XX not form part of the printed specification, but was obtained in
 XX electronic format directly from USPTO at
 XX seqdata.uspto.gov/sequence.html?docID=20020081668
 XX XX
 XX Sequence 328 BP; 97 A; 74 C; 99 G; 56 T; 0 U; 2 Other;

Query Match 34.5%; Score 30; DB 6; Length 328;
 Best Local Similarity 59.3%; Pred. No. 0.63;
 Matches 51; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 AGAGAACCATCTATATGAGAGATTCTGGCTGAGACTTGACGAGGAGCAAGA 60
 DB 103 AGAGAACCATCTATATGAGAGATTCTGGCTGAGACTTGACGAGGAGCAAGA 162
 QY 61 AGAGACTTGACACCACTATGACA 86
 DB 163 AACGACCAAGTAAAGATTGGGAAA 188

RESULT 12
 ID AA216628/c
 ID AA216628 standard; CDNA; 762 BP.
 AC AA216628;
 DT 12-OCT-1999 (first entry)
 DE Human gene expression product CDNA sequence SEQ ID NO:4098.
 KW Human; gene; gene expression product; diagnosis; therapy; probe;
 KW detection; mapping; tissue typing; profiling; forensic; cancer;
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
 OS Homo sapiens.
 PN WO9338972-A2.
 PD 05-AUG-1999.
 PF 28-JAN-1999; 99WC-US001619.
 PR 28-JAN-1999; 98US-0072810P.
 PR 24-FEB-1998; 98US-007354P.
 PR 31-MAR-1998; 98US-0080114P.
 PR 03-APR-1998; 98US-0080515P.
 PR 03-APR-1998; 98US-0080666P.
 PR 21-OCT-1998; 98US-0105234P.
 PR 28-OCT-1998; 98US-0105677P.
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 PI Reinhard C, Glese K, Randazzo F, Kennedy GC, Pot D, Kaseam A;
 PI Lamson G, Drmanac R, Ckvenjakov R, Dickson M, Drmanac S, Labat I;
 PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Grain B;
 XX WPI; 1999-494092/41.
 XX Novel human genes and their expression products which are differentially
 XX expressed in different cell types.
 XX Claim 1; Page 1942; 2479pp; English.

XX The present invention describes a library of human polynucleotides
 XX comprising the sequences given in AA21532 to AA21779. Also described is
 XX a method of detecting differentially expressed genes correlated with the
 XX cancerous state of a mammalian cell, comprising detecting at least one
 XX differentially expressed gene product in a test sample from a cell
 XX suspected of being cancerous, where the gene product is encoded by one of
 XX the 5248 polynucleotide sequences given in AA21532 to AA21779. The
 XX polynucleotides can be used as a source of primers and probes, which can
 XX be used for a variety of purpose, e.g. detection of expression levels,
 XX mapping, tissue typing or profiling, forensics, genetic analysis and
 XX detection of polymorphisms. Polypeptides encoded by the polynucleotides
 XX can be used for raising antibodies for experimental, diagnostic and
 XX therapeutic purposes. The polynucleotides may also be used to construct
 XX arrays for diagnostics (which may be used to determine function of an
 XX encoded protein); and to detect differences in expression levels between
 XX two cells (e.g. to identify abnormal or diseased tissue in a human, to
 XX identify a genetic predisposition or susceptibility to a disease such as
 XX cancer). The polynucleotides of the invention are especially used in the
 XX diagnosis, prognosis and management of colorectal cancer, breast cancer,
 XX and lung cancer. The polynucleotides can also be used to screen for
 XX peptide analogues and antagonists

QY 762 BP; 212 A; 147 C; 163 G; 213 T; 0 U; 27 Other;
 Query Match 33.1%; Score 28.8; DB 2; Length 762;
 Best Local Similarity 60.8%; Pred. No. 2.3;
 Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

| | | |
|----|--|-----|
| OY | GAACCAATCATTAATTAAGGAGATTTTCTGACCTGAGAATTGAGAGGACAAGAGA | 63 |
| Db | 409 GAAAGGCGGTGGNNAACACACACATTTTCTAGCTTCAGAAAGTCAGAGACCCAGATCA | 350 |
| OY | 64 CACTCTGACACCA | 77 |
| Db | 349 CAGCCTGAAACATCA | 336 |

RESULT 13
AAV33481/c
ID AAV33481 standard; DNA; 958 BP.

Human disease related nucleotide kinase-1 (DRNK-1) DNA sequence

Human disease related nucleotide kinase-1; DRNK-1; deoxyguanosine kinase
p21ras; cell proliferation; oncogenesis; cancer; gene therapy;
immune disorder; neurological dysfunction; ss.

Homo sapiens.

| Key | Location/Qualifiers |
|-----|---------------------|
|-----|---------------------|

/*tag=

transit_peptide 61. .135

/note= "Mitochondrial localisation signal"

US5817482-A.

06-OCT-1998.

20-JUN-1997; 97US-00879561.

20-JUN-1997; 9705-00879561.

(INCY-) INCYTE PHARM INC

Hillman JL, Corley NC, Hawkins PR, Guegler KU, Bandman O,

WPI: 1998-556388/4/
D-PSDB: AAW70493

Nucleic acids encoding de:oxy:guanosine kinase - useful for recombinant production of the enzyme for treating diseases caused by lack of the enzyme e.g. cancers caused through loss of enzyme function.

Claim 4; Fig 1A-1C; 53pp; English.

The present sequence represents a human disease related nucleotide kinase CC-1 (DNKR-1) DNA sequence first identified in Invivo Clone 56821 from the CC fibroblast cDNA library (FIBNOT 01). The DNKR-1 DNA sequence is useful CC for the production of the corresponding recombinant enzyme. The invention CC provides DNKR enzymes, which are deoxyguanosine kinases, which catalyze CC the transfer of a terminal phosphate from adenosine triphosphate (ATP) or CC guanine triphosphate (GTP) to guanosine or guandine in the regulation of CC cellular levels of GMP and its corresponding nucleoside triphosphate. As CC GMP levels are known to control the activity of certain oncogenic CC proteins e.g. p21ras, a protein involved in cell proliferation and CC oncogenesis, suppression of the enzyme activity causes high ratios of CC GMP:GDP, promoting oncogenesis. Therefore, diseases (e.g. cancers, immune CC disorders and neurological dysfunction) related to this lack of activity CC may be prevented or treated with the recombinant enzyme, or by gene CC therapy based strategies. Anti-sense constructs of the DNKR encoding CC nucleic acids may also be used for inhibition of over-expression of the CC enzyme

50 Sequence 958 BP; 243 A; 254 C; 228 G; 231 T; 0 U; 2 Other;

| | | | | |
|--------------------------|-------|----------------|-----------|------------|
| Query Match | 32.4% | Score 28.2 | DB 2 | Length 958 |
| Best Local Similarity | 60.8% | Pred. No. 4.2 | | |
| Matches 45; Conservative | 0; | Mismatches 29; | Indels 0; | Gaps 0; |

Oy 4 GAACCATCATTTAATTGAAGTGAGATTTTTCTGGCCCGAGACTTGCAGGGGCAAGAAGA 63
| | | | | | | | | | | | | | | | | | | | | |
Db 929 GAAAGNTGCTGGGAGACACAACATTTTCTCAGCTTCAGAAAGTCAGGAGGCCAGATCA 870

QY 64 CACTCTGGACACCA 77
|||
Db 869 CAGCCTGAACATCA 856

```

RESULT 14
ADD48661
ID      ADD48661 standard; DNA; 2716 BP.

```

DT 29-JAN-2004 (first entry)

DE Rat gene L43592, SEQ ID NO 14367.

KM Rat; ds; gene; pain; neuronal tissue; gene therapy;
KM spinal segmental nerve injury; chronic constriction injury; CCI;
KM spared nerve injury; SNI; Chung.

OS *Rattus norvegicus*.

PN WO2003016475-A2

PD 27-FEB-2003

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P

PR 26-NOV-2001; 2001US-0333347P

PA (GEHO) GEN HOSPITAL CORP

| | |
|------|------------------------------------|
| XX | |
| XXI | D'Amico D Boffetta V Costigliaro M |
| XXII | |

| | |
|----|---------------|
| XX | 2003-2003/200 |
| XX | 2003-2003/200 |

DR GENBANK; L43592

PT New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two, or more isolated RNA
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating

pain (e.g. spinal segmental nerve injury (Chung)), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat DNA (shown in Table 2 of the
 CC specification) which encodes one of the polypeptides of the invention
 CC which is differentially expressed during pain. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 2716 BP; 688 A; 654 C; 657 G; 717 T; 0 U; 0 Other;

CC Query Match 32.0%; Score 27.8; DB 9; Length 2716;

CC Best Local Similarity 62.0%; Pred. No. 8.3;

CC Matches 44; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

CC 17 TTGAAGTGAAGATTTTCTGGGCTGAGATTCAGAGGAGGCAAGACACTCTGGACACC 76

CC 2023 TTGCTGTAGTGCAGTTGTGGGGTGTAGGCTGGCAGAGGAGGAGGAGGCTCCTGGTACT 2082

CC 77 ACTATGGACAG 87

CC 2083 ACTCTGGCCTG 2093

CC RESULT 15

CC ADD32481

CC ID ADD32481 standard; DNA; 444 BP.

CC AC ADD32481;

CC DT 15-JAN-2004 (first entry)

CC DE Human mitochondrial DNA sequence SEQ ID NO:251.

CC DE ds; human; array; mitochondrial; hybridisation; energy-metabolism;

CC KM mitochondrial disease; oxidative phosphorylation dysfunction;

CC KM oxidative stress; apoptosis; aging.

CC OS Homo sapiens.

CC PN WO2003020220-A2.

CC PD 13-MAR-2003.

CC PF 30-AUG-2002; 2002WO-US027886.

CC PR 30-AUG-2001; 2001US-0316323P.

CC PR 31-AUG-2001; 2001CA-02356540.

CC PA (UYEM-) UNTV EMORY.

CC PI Wallace DC, Levy S, Kerstann K, Procaccio V;

CC DR WPI; 2003-300821/29.

CC PT Array containing probes for genes involved in mitochondrial biology,

CC PT useful for determining mitochondrial biology gene expression profiles for

CC PT use in diagnosing pathologies and identifying biochemical pathways.

CC PS Claim 2, SEQ ID NO 251; 201PP; English.

CC The invention relates to a novel array comprising at least two isolated
 CC nucleotide molecules, each molecule having a sequence capable of uniquely
 CC hybridising to a nucleic acid molecule which is an expression product of
 CC a gene involved in mitochondrial biology. The array comprises two or more
 CC isolated nucleic acid molecules or spots, each molecule having a sequence
 CC chosen from sequence of 994 human probes and 2046 mouse probes. An array
 CC of the invention is useful for determining an expression profile of a
 CC mouse or human sample containing nucleic acid, by contacting the array
 CC with the sample under conditions allowing selective hybridisation, and
 CC measuring hybridisation of nucleic acid in the sample to the array to
 CC produce an expression profile. The array is also useful for determining
 CC an expression profile of a first labelled sample containing nucleic acid

CC relative to a second, differently labelled sample containing nucleic
 CC acid. The second sample is a reference or a standard. An array is useful
 CC for determining an expression profile diagnostic of an energy-metabolism-
 CC related physiological condition. An array of the invention is useful for
 CC determining mitochondrial biology gene expression profiles of organisms,
 CC such as human, mice and closely related species, tissue and organs of
 CC such organisms, which are useful for determining expression profiles
 CC diagnostic of energy metabolism-related physiological conditions,
 CC diagnosing such physiological conditions, identifying biochemical
 CC pathways, genes, and mutations involved in such physiological conditions,
 CC identifying therapeutic agents useful for preventing and/or treating such
 CC physiological conditions, evaluating and/or monitoring the efficacy of
 CC such therapies, and creating and identifying animal models of human
 CC energy metabolism-related physiological conditions. An array is also
 CC useful for defining expression signatures or profiles for mitochondrial
 CC diseases, as well as distinguishing clinical disorders that result from
 CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
 CC apoptosis and aging. An array of the invention contains probes of genes
 CC not previously recognised to participate in mitochondrial biology. The
 CC sequences shown in ADD32231-ADD33223 represent human mitochondrial DNA
 CC clones used to make the probes of the invention. Some sequences are not
 CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
 CC 1906, 2408 and 2643.

CC SQ Sequence 444 BP; 128 A; 95 C; 106 G; 114 T; 0 U; 1 Other;

CC Query Match 31.7%; Score 27.6; DB 9; Length 444;

CC Best Local Similarity 60.8%; Pred. No. 5.4; 29; Indels 0; Gaps 0;

CC Matches 45; Conservative 0; Mismatches 29;

CC 4 GAACATCATTAATTAAGTGAATTTTCTGGCCTGAGACTTGGAGGAGGCAAGAAGA 63

CC DB 129 GAAGGTGGTTGGAGACACAAATTTTCTAGCTTCAGAAAGTAGGAGGCCAGATCA 188

CC 64 CACTTGGACACCA 77

CC DB 189 CAGCCTGAACATCA 202

Search completed: March 12, 2004, 18:08:05
 Job time : 35.8377 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 17:39:25 ; Search time 6.47321 Seconds
(without alignments)
7458.540 Million cell updates/sec

Title: US-09-966-880A-11

Perfect score: 87

Sequence: 1 agaaacacacataatga.....ctggacacacacatgacag 87

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| C 1 | 28.2 | 32.4 | 958 | 1 | US-08-879-561-2 |
| C 2 | 27.6 | 31.7 | 587 | 4 | US-09-833-381-916 |
| C 3 | 27.6 | 31.7 | 843 | 2 | US-08-458-657-1 |
| C 4 | 27.6 | 31.7 | 843 | 3 | US-08-860-995-1 |
| C 5 | 27.6 | 31.7 | 843 | 5 | PCR-US95-00532A-1 |
| C 6 | 27.6 | 31.7 | 1003 | 4 | US-09-833-381-917 |
| C 7 | 25.2 | 29.0 | 5448 | 4 | US-09-620-312D-246 |
| C 8 | 25.2 | 29.0 | 5532 | 4 | US-09-620-312D-245 |
| C 9 | 25.2 | 29.0 | 5532 | 4 | US-09-545-481-3 |
| C 10 | 24.6 | 28.3 | 786431 | 4 | US-09-751-389-3 |
| C 11 | 24.4 | 28.0 | 492 | 3 | US-08-998-416-41 |
| C 12 | 24.4 | 28.0 | 1631 | 3 | US-08-895-590-43 |
| C 13 | 24.2 | 27.8 | 2431 | 4 | US-09-023-655-810 |
| C 14 | 24.2 | 27.8 | 2546 | 4 | US-09-598-401C-95 |
| C 15 | 24.2 | 27.8 | 3016 | 4 | US-09-016-434-1110 |
| C 16 | 24.2 | 27.8 | 3035 | 5 | PCR-US94-09235-1 |
| C 17 | 24.2 | 27.8 | 168575 | 4 | US-09-426-290-1 |
| C 18 | 24.2 | 27.6 | 546 | 4 | US-09-621-976-2448 |
| C 19 | 23.8 | 27.4 | 1555 | 4 | US-09-227-357-25 |
| C 20 | 23.8 | 27.4 | 252 | 3 | US-09-130-616-176 |
| C 21 | 23.8 | 27.4 | 4892 | 4 | US-09-620-312D-1065 |
| C 22 | 23.8 | 27.4 | 5537 | 1 | US-08-135-511-32 |
| C 23 | 23.8 | 27.4 | 5537 | 1 | US-08-483-852-9 |
| C 24 | 23.8 | 27.4 | 5537 | 1 | US-08-361-458-4 |
| C 25 | 23.8 | 27.4 | 5537 | 1 | US-08-477-953-9 |
| C 26 | 23.8 | 27.4 | 5537 | 1 | US-08-187-453-32 |
| C 27 | 23.8 | 27.4 | 5537 | 1 | US-08-562-985A-5 |

| | | | | | | |
|------|------|------|--------|---|-------------------|--------------------|
| 28 | 23.8 | 27.4 | 5537 | 2 | US-08-477-952-9 | Sequence 9, Appli |
| C 29 | 23.8 | 27.4 | 9573 | 4 | US-09-220-132-168 | Sequence 168, App |
| C 30 | 23.8 | 27.4 | 51952 | 3 | US-08-947-823-1 | Sequence 1, Appli |
| C 31 | 23.8 | 27.4 | 64467 | 4 | US-09-803-671B-3 | Sequence 3, Appli |
| C 32 | 23.6 | 27.1 | 128779 | 4 | US-09-497-855A-38 | Sequence 38, Appli |
| C 33 | 23.4 | 26.9 | 551 | 3 | US-09-040-984-29 | Sequence 29, Appli |
| C 34 | 23.4 | 26.9 | 551 | 4 | US-09-123-912-29 | Sequence 29, Appli |
| C 35 | 23.4 | 26.9 | 551 | 4 | US-09-643-597-29 | Sequence 29, Appli |
| C 36 | 23.4 | 26.9 | 551 | 4 | US-09-480-884A-29 | Sequence 29, Appli |
| C 37 | 23.4 | 26.9 | 551 | 4 | US-09-542-615A-29 | Sequence 29, Appli |
| C 38 | 23.4 | 26.9 | 551 | 4 | US-09-606-421B-29 | Sequence 29, Appli |
| C 39 | 23.4 | 26.9 | 551 | 4 | US-09-221-107-29 | Sequence 29, Appli |
| C 40 | 23.4 | 26.9 | 4693 | 4 | US-09-462-561B-8 | Sequence 8, Appli |
| C 41 | 23.4 | 26.9 | 269223 | 4 | US-09-596-002-41 | Sequence 41, Appli |
| C 42 | 23.2 | 26.7 | 660 | 3 | US-09-248-528-14 | Sequence 14, Appli |
| C 43 | 23.2 | 26.7 | 660 | 3 | US-09-549-108-14 | Sequence 14, Appli |
| C 44 | 23.2 | 26.7 | 660 | 3 | US-09-549-111-14 | Sequence 14, Appli |
| C 45 | 23.2 | 26.7 | 660 | 3 | US-09-549-106-14 | Sequence 14, Appli |

ALIGNMENTS

RESULT 1
US-08-879-561-2/c
Sequence 2, Application US/08879561
Patent No. 5817482
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Hawkins, Phillip R.
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/879,561
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0325 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: FIBRONOT1
CLONE: 56821
US-08-879-561-2

Query Match 32.4%; Score 28.2; DB 1; Length 958;
Best Local Similarity 60.8%; Pred. No. 0.44; 29; Indels 0; Gaps 0;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 4 GAACCATCATTAATTAAGTGAATTTTCTGGCCTGAGACTTGCAAGGAGGCAAGAAGA 63
DB 929 GAAAGTGTGGGAGACACACATTTTCTAGCTTCAAGAAAGTCAGGAGCCCAAGATCA 870

QY 64 CACTGTGACACCA 77
DB 869 CAGCTGACATCA 856

RESULT 2
US-09-833-381-916
Sequence 916; Application US/09833381
Patent No. 6672186

GENERAL INFORMATION:
APPLICANT: ROBINSON, Keith E.
TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
FILE REFERENCE: 5800-119

CURRENT APPLICATION NUMBER: US/09/833,381
PRIOR FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 09/516,448
NUMBER OF SEQ ID NOS: 2050
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 916
LENGTH: 587
TYPE: DNA
ORGANISM: Homo sapiens
US-09-833-381-916

Query Match 31.7%; Score 27.6; DB 4; Length 587;
Best Local Similarity 60.8%; Pred. No. 0.62; 29; Indels 0; Gaps 0;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 4 GAACCATCATTAATTAAGTGAATTTTCTGGCCTGAGACTTGCAAGGAGGCAAGAAGA 63
DB 132 GAAAGTGTGGGAGACACACATTTTCTAGCTTCAAGAAAGTCAGGAGCCCAAGATCA 191

QY 64 CACTGTGACACCA 77
DB 192 CAGCTGACATCA 205

RESULT 3
US-08-458-657-1/c
Sequence 1; Application US/08458657
Patent No. 5914258
GENERAL INFORMATION:
APPLICANT: WEI, ET AL.
TITLE OF INVENTION: Human Deoxycytidine Kinase 2
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,657
FILING DATE: June 2, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00532

FILING DATE: 13 JAN 95
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-371
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 BASE PAIRS
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
US-08-458-657-1

Query Match 31.7%; Score 27.6; DB 2; Length 843;
Best Local Similarity 60.8%; Pred. No. 0.7; 29; Indels 0; Gaps 0;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 4 GAACCATCATTAATTAAGTGAATTTTCTGGCCTGAGACTTGCAAGGAGGCAAGAAGA 63
DB 690 GAAAGTGTGGGAGACACACATTTTCTAGCTTCAAGAAAGTCAGGAGCCCAAGATCA 631

QY 64 CACTGTGACACCA 77
DB 630 CAGCTGACATCA 617

RESULT 4
US-08-860-995-1/c
Sequence 1; Application US/08860995
Patent No. 6063376
GENERAL INFORMATION:
APPLICANT: WEI, ET AL.
TITLE OF INVENTION: Human Deoxycytidine Kinase 2
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/860,995
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00532
FILING DATE: 13 JAN 95
ATTORNEY/AGENT INFORMATION:
NAME: PERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-246
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 BASE PAIRS
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA

US-08-860-995-1

Query Match 31.7%; Score 27.6; DB 3; Length 843;
Best Local Similarity 60.8%; Pred. No. 0.7;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 4 GAACCATCATTAATTGAAGTGAAGATTTTCTGCGCTGAGACTTTCAGAGGAGCAAGAGA 63
DB 690 GAAAGGTGGTTGGAGAGACACAACTTTTCTAGCTTCAAGAAAGTCAGGAGGCCCAATCA 631

QY 64 CACTCTGACACCA 77
DB 630 CAGCCTGAACATCA 617

RESULT 5

PCT-US95-00532A-1/c
Sequence 1, Application PC/TUS9500532A

GENERAL INFORMATION:

APPLICANT: WEI, ET AL.

TITLE OF INVENTION: Human Deoxycytidine Kinase 2

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESSES:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

ADDRESSEE: CECCHI, STEWART & OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

FILING DATE: 13 JAN 95

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: FERRARO, GREGORY D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-246

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 843 BASE PAIRS

TYPE: NUCLEIC ACID

STRANDEDNESS: SINGLE

TOPOLOGY: LINEAR

MOLECULE TYPE: CDNA

PCT-US95-00532A-1

Query Match 31.7%; Score 27.6; DB 5; Length 843;

Best Local Similarity 60.8%; Pred. No. 0.7;

Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 4 GAACCATCATTAATTGAAGTGAAGATTTTCTGCGCTGAGACTTTCAGAGGAGCAAGAGA 63

DB 690 GAAAGGTGGTTGGAGAGACACAACTTTTCTAGCTTCAAGAAAGTCAGGAGGCCCAATCA 631

QY 64 CACTCTGACACCA 77

DB 630 CAGCCTGAACATCA 617

DB 630 CAGCCTGAACATCA 617

DB 630 CAGCCTGAACATCA 617

DB 630 CAGCCTGAACATCA 617

DB 630 CAGCCTGAACATCA 617

DB 630 CAGCCTGAACATCA 617

DB 630 CAGCCTGAACATCA 617

US-09-833-381-917

Sequence 917, Application US/09833381
Patent No. 6672186
GENERAL INFORMATION:

APPLICANT: ROBISON, Keith E.

TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs

FILE REFERENCE: 5800-119

CURRENT APPLICATION NUMBER: US/09/833,381

CURRENT FILING DATE: 2001-04-11

PRIOR APPLICATION NUMBER: 09/516,448

PRIOR FILING DATE: 2000-02-29

NUMBER OF SEQ ID NOS: 2050

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 917

LENGTH: 1003

TYPE: DNA

ORGANISM: Homo sapiens

US-09-833-381-917

Query Match 31.7%; Score 27.6; DB 4; Length 1003;

Best Local Similarity 60.8%; Pred. No. 0.75;

Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 4 GAACCATCATTAATTGAAGTGAAGATTTTCTGCGCTGAGACTTTCAGAGGAGCAAGAGA 63

DB 178 GAAAGGTGGTTGGAGAGACACAACTTTTCTAGCTTCAAGAAAGTCAGGAGGCCCAATCA 237

QY 64 CACTCTGACACCA 77

DB 238 CAGCCTGAACATCA 251

DB 238 CAGCCTGAACATCA 251

RESULT 7

US-09-620-312D-246

Sequence 246, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyun

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aldong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yungling

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinbacht

APPLICANT: Dmanac, Radose T

TITLE OF INVENTION: No. 6569662el Nucleic Acids and

FILE REFERENCE: 784C1P2B

CURRENT APPLICATION NUMBER: US/09/620,312D

CURRENT FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pt_FL_genes Version 1.0

SEQ ID NO 246

LENGTH: 5448

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (534)..(4178)

US-09-620-312D-246

Query Match 29.0%; Score 25.2; DB 4; Length 5448;
Best Local Similarity 62.9%; Pred. No. 11;
Matches 39; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 AGAGAACCATCATTAATTTGAAGTGAATTTTCTGGCCTGAGACTGAGGAGGAGCAAGA 60
DB 1910 AGATAATGAGCTAATGGAAGAAAGAAATATATGCTGAGAGCTTCCAGAAAGCAAAA 1969

QY 61 AG 62
DB 1970 AG 1971

RESULT 8
US-09-620-312D-245
; Sequence 245, Application US/09620312D

; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonshong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhilwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radote T
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_Fl_genes Version 1.0
; SEQ ID NO 245
; LENGTH: 5532
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (534)..(4262)
US-09-620-312D-245

Query Match 29.0%; Score 25.2; DB 4; Length 5532;
Best Local Similarity 62.9%; Pred. No. 11;
Matches 39; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 AGAGAACCATCATTAATTTGAAGTGAATTTTCTGGCCTGAGACTGAGGAGGAGCAAGA 60
DB 1910 AGATAATGAGCTAATGGAAGAAAGAAATATATGCTGAGAGCTTCCAGAAAGCAAAA 1969

QY 61 AG 62
DB 1970 AG 1971

RESULT 9
US-09-545-481-3
; Sequence 3, Application US/09545481
; Patent No. 6451319

; GENERAL INFORMATION:
; APPLICANT: Chiang, H.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: No. 6451319el Recombinant And Mutant Adenoviruses
; FILE REFERENCE: SY0993K US
; CURRENT APPLICATION NUMBER: US/09/545,481
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,766
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 34185
; TYPE: DNA
; ORGANISM: Bovine adenovirus type 1
US-09-545-481-3

Query Match 28.7%; Score 25; DB 4; Length 34185;
Best Local Similarity 56.8%; Pred. No. 25;
Matches 46; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 6 ACCATCATTAATTTGAAGTGAATTTTCTGGCCTGAGACTGAGGAGGAGCAAGAACA 65
DB 29304 AACATCTTTATTTGAGCTGGAATTTTGTCATGACTCTTAAGATTCACCAACAC 29363
QY 66 CTCGTGACACCACTATGACA 86
DB 29364 CACTCGATGCAACCTTATACA 29384

RESULT 10
US-09-751-389-3/C
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001067
; CURRENT APPLICATION NUMBER: US/09/751,389
; CURRENT FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(786431)
; OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3

Query Match 28.3%; Score 24.6; DB 4; Length 786431;
Best Local Similarity 59.2%; Pred. No. 1e+02;
Matches 42; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 15 AATTGAAGTGAATTTTCTGGCCTGAGACTTGCAGGAGGAGCAAGACACTTGTGACA 74
DB 617689 AGTTTAAAGAAATGCTTGGATTAGTTTCTCAGAGATGAAATACATCTTGGCA 617630
QY 75 CCACTATGAC 85
DB 617629 GCACTATCTAC 617619

RESULT 11
US-08-998-416-41/c
; Sequence 41, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter

APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reibischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSEYP11
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Weigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGCL976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 492 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE: PAGO10UP
ORGANISM: PAGO10UP
US-08-998-416-41
Query Match 28.0%; Score 24.4; DB 3; Length 492;
Best Local Similarity 68.0%; Pred. No. 9.1; Indels 16; Gaps 0;
Matches 34; Conservative 0; Mismatches 16; Indels 16; Gaps 0;
QY 21 AGTGAATTTTCTGCGCTGAGACTTGACGAGGAGGCAAGACACTCTG 70
DB 461 ATTGAATTTTCTGCTGCTGAGGCTTCACTGAGCTGGGATGAAGGCTCG 412
RESULT 12
US-08-895-590-43/C
Sequence 43, Application US/08895590
Patent No. 6207410
GENERAL INFORMATION:
APPLICANT: Hall, Linda M.
APPLICANT: Ren, Dejian
APPLICANT: Zheng, Wei
APPLICANT: Dubald, Manuel Marcel Paul
TITLE OF INVENTION: Genes Encoding an Insect Calcium Channel
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314-3187

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/895,590
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/374,888
FILING DATE: 19-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm M.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 022650-263
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1631 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 33..1631
US-08-895-590-43
Query Match 28.0%; Score 24.4; DB 3; Length 1631;
Best Local Similarity 60.6%; Pred. No. 14; Indels 26; Gaps 0;
Matches 40; Conservative 0; Mismatches 26; Indels 26; Gaps 0;
QY 15 AATTGAAGTGAATTTTCTGCGCTGAGACTTGACGAGGAGGCAAGACACTCTGACA 74
DB 1084 ATTGAAGGAGAGGTTTTCGCTCGACTTCCGACGAGTGTGATCAAGCTGTAA 1025
QY 75 CCACTA 80
DB 1024 ACCTTA 1019
RESULT 13
US-09-023-655-810/C
Sequence 810, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Sellhammer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
TITLE OF INVENTION: EXPRESSION
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 810:
SEQUENCE CHARACTERISTICS:
LENGTH: 2431 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAIT102
CLONE: 757560
US-09-023-655-810

Query Match 27.8%; Score 24.2; DB 4; Length 2431;
Best Local Similarity 66.0%; Pred. No. 19;
Matches 35; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 8 CATCATTAATGAGTGAATTTCTGGCCTGAGACTTGCGAGGAGCAAGA 60
DB 616 CATCATTAATGAGTGAATTTCTGGCCTGAGACTTGCGAGGAGCAAGA 564

RESULT 14
US-09-598-401C-95/C
Sequence 95, Application US/09598401C
Patent No. 6596925
GENERAL INFORMATION:
APPLICANT: Perera, J. Ranjan
APPLICANT: Eagleton, Clare
APPLICANT: Rice, Stephen J.
TITLE OF INVENTION: Compositions and Methods for the
FILE OF INVENTION: Modification of Gene Expression
FILE REFERENCE: 11000.1036C2
CURRENT APPLICATION NUMBER: US/09/598,401C
CURRENT FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: U.S. No. 6596925 60/146,591
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: PCT/NZ00/00018
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: U.S. No. 6596925 09/276,599
PRIOR FILING DATE: 1999-03-25
NUMBER OF SEQ ID NOS: 120
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 95
LENGTH: 2546
TYPE: DNA
ORGANISM: Pinus radiata
US-09-598-401C-95

Query Match 27.8%; Score 24.2; DB 4; Length 2546;
Best Local Similarity 57.1%; Pred. No. 20;
Matches 44; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 5 AACCATTAATGAGTGAATTTCTGGCCTGAGACTTGCGAGGAGCAAGA 64
DB 2472 AACCATTAATGAGTGAATTTCTGGCCTGAGACTTGCGAGGAGCAAGA 2413

QY 65 ACTTGAGACCACTAT 81
DB 2412 ACTTGAGACCACTAT 2396

RESULT 15
US-09-016-434-1110
Sequence 1110, Application US/09016434
Patent No. 6500938

GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
APPLICATION DATA:
PRIOR APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1110:
SEQUENCE CHARACTERISTICS:
LENGTH: 3016 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 91321593
US-09-016-434-1110

Query Match 27.8%; Score 24.2; DB 4; Length 3016;
Best Local Similarity 66.0%; Pred. No. 21;
Matches 35; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 AGAGAACCATTAATGAGTGAATTTCTGGCCTGAGACTTGCGAGGAG 53
DB 1294 AGAGAACCATTAATGAGTGAATTTCTGGCCTGAGACTTGCGAGGAG 1346

Search completed: March 13, 2004, 00:51:06
Job time : 8.47321 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 21:22:51 / Search time 30.53 Seconds

(Without alignments)
10491.276 Million cell updates/sec

Title: US-09-966-880A-11

Perfect score: 87

Sequence: 1 agagaaccatcatatga.....ctggacacatcatgacag 87

Scoring table:

IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2432557 seqs, 1840798884 residues

465114

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1 | 87 | 100.0 | 87 | US-09-966-880A-11 | Sequence 11, Appl |
| 2 | 87 | 100.0 | 2818 | US-09-966-880A-7 | Sequence 7, Appl |
| 3 | 87 | 100.0 | 5514 | US-09-966-880A-9 | Sequence 9, Appl |
| 4 | 87 | 100.0 | 11204 | US-09-966-880A-35 | Sequence 35, Appl |
| 5 | 41 | 47.1 | 2440 | US-09-966-880A-1 | Sequence 1, Appl |
| 6 | 30 | 34.5 | 328 | US-09-728-446-1014 | Sequence 1014, Ap |
| 7 | 28 | 32.2 | 809 | US-10-027-632-164874 | Sequence 164874, |
| 8 | 27.6 | 31.7 | 327 | US-10-085-783A-7046 | Sequence 7046, Ap |
| 9 | 27.6 | 31.7 | 327 | US-10-242-535A-7046 | Sequence 7046, Ap |
| 10 | 27.6 | 31.7 | 333 | US-09-796-692-5294 | Sequence 5294, Ap |
| 11 | 27.6 | 31.7 | 333 | US-10-040-862-5294 | Sequence 5294, Ap |
| 12 | 27.6 | 31.7 | 333 | US-10-057-475B-5294 | Sequence 5294, Ap |
| 13 | 27.6 | 31.7 | 333 | US-10-154-864B-5294 | Sequence 5294, Ap |
| 14 | 27.6 | 31.7 | 587 | US-09-833-381-916 | Sequence 916, App |
| 15 | 27.6 | 31.7 | 1003 | US-09-833-381-917 | Sequence 917, App |

| | | | | | | |
|----|------|------|--------|----|----------------------|--------------------|
| 16 | 27.4 | 31.5 | 947 | 10 | US-09-974-879-101 | Sequence 101, App |
| 17 | 27.4 | 31.5 | 947 | 10 | US-09-305-736-101 | Sequence 101, App |
| 18 | 27.4 | 31.5 | 947 | 11 | US-09-818-683-101 | Sequence 101, App |
| 19 | 27.4 | 31.5 | 947 | 12 | US-10-621-401-101 | Sequence 101, App |
| 20 | 27.2 | 31.3 | 1119 | 14 | US-10-158-846-1354 | Sequence 1354, A |
| 21 | 26.8 | 30.8 | 130320 | 15 | US-10-408-168-1 | Sequence 1, Appl |
| 22 | 26.6 | 30.6 | 1706 | 12 | US-10-424-599-44672 | Sequence 44672, A |
| 23 | 26.4 | 30.3 | 593 | 15 | US-10-027-632-253079 | Sequence 253079, |
| 24 | 26.4 | 30.3 | 593 | 15 | US-10-027-632-253080 | Sequence 253080, |
| 25 | 26.4 | 30.3 | 593 | 15 | US-10-027-632-253081 | Sequence 253081, |
| 26 | 26.4 | 30.3 | 609 | 15 | US-10-027-632-10277 | Sequence 10277, A |
| 27 | 26.4 | 30.3 | 1101 | 15 | US-10-027-632-11144 | Sequence 31144, A |
| 28 | 26.4 | 30.3 | 1164 | 15 | US-10-027-632-10459 | Sequence 30459, A |
| 29 | 26.2 | 30.1 | 630 | 15 | US-10-027-632-96683 | Sequence 96683, A |
| 30 | 26.2 | 30.1 | 630 | 15 | US-10-027-632-108335 | Sequence 308335, |
| 31 | 26.2 | 30.1 | 814 | 15 | US-10-027-632-173408 | Sequence 173408, |
| 32 | 26.2 | 30.1 | 32154 | 9 | US-09-764-877-3433 | Sequence 3433, Ap |
| 33 | 26.2 | 30.1 | 32154 | 9 | US-10-242-515-3433 | Sequence 3433, Ap |
| 34 | 26.2 | 30.1 | 32154 | 15 | US-10-101-510-703 | Sequence 703, Appl |
| 35 | 26 | 29.9 | 26 | 9 | US-09-966-880A-31 | Sequence 31, Appl |
| 36 | 26 | 29.9 | 417 | 14 | US-10-082-828A-62 | Sequence 62, Appl |
| 37 | 26 | 29.9 | 435 | 9 | US-09-796-692-8915 | Sequence 8915, Ap |
| 38 | 26 | 29.9 | 435 | 14 | US-10-040-862-8915 | Sequence 8915, Ap |
| 39 | 26 | 29.9 | 435 | 15 | US-10-057-475B-8915 | Sequence 8915, Ap |
| 40 | 26 | 29.9 | 435 | 15 | US-10-154-864B-8915 | Sequence 8915, Ap |
| 41 | 26 | 29.9 | 580 | 15 | US-10-027-632-279917 | Sequence 279917, A |
| 42 | 26 | 29.9 | 779 | 15 | US-10-027-632-13154 | Sequence 13154, A |
| 43 | 26 | 29.9 | 887 | 15 | US-10-103-959A-1 | Sequence 1, Appl |
| 44 | 26 | 29.9 | 985 | 15 | US-10-103-959A-4 | Sequence 4, Appl |
| 45 | 26 | 29.9 | 1328 | 14 | US-10-082-828A-63 | Sequence 63, Appl |

ALIGNMENTS

RESULT 1

US-09-966-880A-11

Sequence 11, Application US/09966880A

Patent No. US2002016473A1

GENERAL INFORMATION:

APPLICANT: Honjo, Tasuku

TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE

FILE REFERENCE: 06501-088001

CURRENT APPLICATION NUMBER: US/09/966, 880A

PRIOR FILING DATE: 2001-09-28

PRIOR APPLICATION NUMBER: PCT/JP00/01918

PRIOR FILING DATE: 2000-03-28

PRIOR APPLICATION NUMBER: JP 11-371382

PRIOR FILING DATE: 1999-12-27

PRIOR APPLICATION NUMBER: JP 11-178999

PRIOR FILING DATE: 1999-06-24

PRIOR APPLICATION NUMBER: JP 11-87192

PRIOR FILING DATE: 1999-03-29

NUMBER OF SEQ ID NOS: 36

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11

LENGTH: 87

TYPE: DNA

ORGANISM: Homo sapiens

US-09-966-880A-11

Query Match 100.0%; Score 87; DB 9; Length 87;

Best Local Similarity 100.0%; Pred. No. 7.6e-23;

Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAACAATCAATTAATTAAGTGAATTTTCTGGCTTGACCTTGACGAGGACGAGA 60

DB 1 AGAACAATCAATTAATTAAGTGAATTTTCTGGCTTGACCTTGACGAGGACGAGA 60

QY 61 AGAACAATCAATTAATTAAGTGAATTTTCTGGCTTGACCTTGACGAGGACGAGA 60

DB 61 AGAACAATCAATTAATTAAGTGAATTTTCTGGCTTGACCTTGACGAGGACGAGA 60

```
RESULT 2
US-09-966-880A-7
; Sequence 7, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2818
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (80)...(673)
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)...(79)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (677)...(2818)
US-09-966-880A-7

Query Match
Best Local Similarity 100.0%; Score 87; DB 9; Length 2818;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGACCATCATTAATTGAAGTGAATTTTCTGCGCTGAGACTTGACGGAGGCAAGA 60
DB 1 AGAAGACCATCATTAATTGAAGTGAATTTTCTGCGCTGAGACTTGACGGAGGCAAGA 60

QY 61 AGACACTCTGAGACACCACTATGACAG 87
DB 61 AGACACTCTGAGACACCACTATGACAG 87

RESULT 3
US-09-966-880A-9
; Sequence 9, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
```

```
LENGTH: 5514
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Intron
; LOCATION: (1)...(1031)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1032)...(1118)
; NAME/KEY: Intron
; LOCATION: (1119)...(5514)
US-09-966-880A-9

Query Match
Best Local Similarity 100.0%; Score 87; DB 9; Length 5514;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGACCATCATTAATTGAAGTGAATTTTCTGCGCTGAGACTTGACGGAGGCAAGA 60
DB 1032 AGAAGACCATCATTAATTGAAGTGAATTTTCTGCGCTGAGACTTGACGGAGGCAAGA 1091

QY 61 AGACACTCTGAGACACCACTATGACAG 87
DB 1092 AGACACTCTGAGACACCACTATGACAG 1118

RESULT 4
US-09-966-880A-35
; Sequence 35, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 11204
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-966-880A-35

Query Match
Best Local Similarity 100.0%; Score 87; DB 9; Length 11204;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGACCATCATTAATTGAAGTGAATTTTCTGCGCTGAGACTTGACGGAGGCAAGA 60
DB 442 AGAAGACCATCATTAATTGAAGTGAATTTTCTGCGCTGAGACTTGACGGAGGCAAGA 501

QY 61 AGACACTCTGAGACACCACTATGACAG 87
DB 502 AGACACTCTGAGACACCACTATGACAG 528

RESULT 5
US-09-966-880A-1
; Sequence 1, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
```

```
APPLICANT: Muramatsu, Masamichi
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966,880A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2440
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (93)...(686)
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)...(92)
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (690)...(2440)
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(2440)
OTHER INFORMATION: n = A,T,C or G
US-09-966-880A-1
```

```
Query Match
Best Local Similarity 47.1%; Score 41; DB 9; Length 2440;
Matches 64; Conservative 0; Mismatches 20; Indels 1; Gaps 1;
```

```
QY 4 GAACCATCATTAATTAAGTGAATTTTCTGGCTGAGACTTGAGGGGAGCAAG-AAG 62
DB 16 GAAGAGCGCTTCTTGAAGCAAGCTTCTTGGCTTAAGACTTGAAGGAGTCAAGAAAG 75
QY 63 AACTCTGACACCCACTATGACAG 87
DB 76 TCACGCTGAGACCGCATATGACAG 100
```

```
RESULT 6
US-09-728-446-1014
Sequence 1014, Application US/09728446
Patent No. US20020081668A1
GENERAL INFORMATION:
APPLICANT: Friedrich, Glenn
APPLICANT: Zambrowicz, Brian
TITLE OF INVENTION: No. US20020081668A1 Murine Polynucleotide Sequences
TITLE OF INVENTION: and Mutant Cells and Mutant Animals Defined Thereby
FILE REFERENCE: LEX-0101-USA
CURRENT APPLICATION NUMBER: US/09/728,446
CURRENT FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/168,270
PRIOR FILING DATE: 1999-12-01
NUMBER OF SEQ ID NOS: 1461
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1014
LENGTH: 328
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(328)
OTHER INFORMATION: n = A,T,C or G
US-09-728-446-1014
```

```
Query Match
Best Local Similarity 34.5%; Score 30; DB 9; Length 328;
Matches 51; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
QY 1 AGAAGACCATTAATTAAGTGAATTTTCTGGCTGAGACTTGAGGGGAGCAAGCA 60
DB 103 AGAAGACCATTAATTAAGTGAATTTTCTGGCTGAGACTTGAGGGGAGCAAGCA 162
QY 61 AGAAGACCATTAATTAAGTGAATTTTCTGGCTGAGACTTGAGGGGAGCAAGCA 86
DB 163 AACGACCAAGTAAGATTCGGGAAA 188
```

```
RESULT 7
US-10-027-632-164874
Sequence 164874, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
```

```
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
POLYMORPHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 164874
LENGTH: 809
TYPE: DNA
ORGANISM: Human
US-10-027-632-164874
```

```
Query Match
Best Local Similarity 32.2%; Score 28; DB 15; Length 809;
Matches 43; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
```

```
QY 10 TCATTAATTAAGTGAATTTTCTGGCTGAGACTTGAGGGGAGCAAGCAAGCACT 69
DB 335 TGAGCACTTAATTAAGTGAATTTTCTGGCTGAGACTTGAGGGGAGCAAGCAAGCACT 394
QY 70 GGAACCA 77
DB 395 GGAATTC 402
```

```
RESULT 8
US-10-085-783A-7046/C
Sequence 7046, Application US/10085783A
Publication No. US20040037841A1
GENERAL INFORMATION:
APPLICANT: ChondroGene Inc.
APPLICANT: Ilew, C.C.
TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
FILE REFERENCE: 4231/2002
CURRENT APPLICATION NUMBER: US/10/085,783A
CURRENT FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/305,340
PRIOR FILING DATE: 2001-07-13
```


PRIOR APPLICATION NUMBER: US 60/275,017
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: US 60/271,955
PRIOR FILING DATE: 2001-02-28
NUMBER OF SEQ ID NOS: 58994
SOFTWARE: Patent version 3.2
SEQ ID NO 7046
LENGTH: 327
TYPE: DNA
ORGANISM: Human
US-10-085-783A-7046

Query Match
Best Local Similarity 60.8%; Pred. No. 3.1;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 4 GAACCATCATTAATTGAGTGAATTTTCTGGCCTGAGACTTGACAGGAGGCAAGAGA 63
DB 189 GAAAGGTGTTGGGAGACACAACATTTTCTAGCTTCGAAAGTCAGGGAGCCGAGATCA 130
QY 64 CACTCTGACACCA 77
DB 129 CAGCCTGAACATCA 116

RESULT 9
US-10-242-535A-7046/C
Sequence 7046, Application US/10242535A
Publication No. US20040013663A1
GENERAL INFORMATION:
APPLICANT: ChondroGene Inc.

TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
FILE REFERENCE: 4231/2005
CURRENT APPLICATION NUMBER: US/10/242,535A
CURRENT FILING DATE: 2002-09-12
PRIOR APPLICATION NUMBER: US 10/085,783
PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/305,340
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/275,017
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: US 60/271,955
PRIOR FILING DATE: 2001-02-28
NUMBER OF SEQ ID NOS: 58994
SOFTWARE: Patent version 3.2
SEQ ID NO 7046
LENGTH: 327
TYPE: DNA
ORGANISM: Human
US-10-242-535A-7046

Query Match
Best Local Similarity 60.8%; Pred. No. 3.1;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 4 GAACCATCATTAATTGAGTGAATTTTCTGGCCTGAGACTTGACAGGAGGCAAGAGA 63
DB 189 GAAAGGTGTTGGGAGACACAACATTTTCTAGCTTCGAAAGTCAGGGAGCCGAGATCA 130
QY 64 CACTCTGACACCA 77
DB 129 CAGCCTGAACATCA 116

RESULT 10
US-09-796-692-5294/C
Sequence 5294, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077,001200
CURRENT APPLICATION NUMBER: US/09/796,692
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5294
LENGTH: 333
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: (9)
OTHER INFORMATION: n=A,T,C or G
US-09-796-692-5294

Query Match
Best Local Similarity 60.8%; Pred. No. 3.1;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 4 GAACCATCATTAATTGAGTGAATTTTCTGGCCTGAGACTTGACAGGAGGCAAGAGA 63
DB 248 GAAAGGTGTTGGGAGACACAACATTTTCTAGCTTCGAAAGTCAGGGAGCCGAGATCA 189
QY 64 CACTCTGACACCA 77
DB 188 CAGCCTGAACATCA 175

RESULT 11
US-10-040-862-5294/C
Sequence 5294, Application US/10040862
Publication No. US20030078396A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,862
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17

```

; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5294
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (9)
; OTHER INFORMATION: n=A,T,C or G
US-10-040-862-5294

Query Match      31.7%; Score 27.6; DB 14; Length 333;
Best Local Similarity 60.8%; Pred. No. 3.1;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY      4  GAACCATCATTAATGAGAGATTTCCTGCGCTGAGACTTGAGGGAGGCAAGAGA 63
DB      248  GAAGGTGGTGGAGACACACATTTCTAGCTTCAGAAAGTCAGGGAGCCAGATCA 189
QY      64  CACTCTGACACCA 77
DB      188  CAGCCTGAACATCA 175

RESULT 12
US-10-057-475B-5294/C
; Sequence 5294, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
```

```

; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5294
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(333)
; OTHER INFORMATION: n = G, A, C or T
US-10-057-475B-5294

Query Match      31.7%; Score 27.6; DB 15; Length 333;
Best Local Similarity 60.8%; Pred. No. 3.1;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY      4  GAACCATCATTAATGAGAGATTTCCTGCGCTGAGACTTGAGGGAGGCAAGAGA 63
DB      248  GAAGGTGGTGGAGACACACATTTCTAGCTTCAGAAAGTCAGGGAGCCAGATCA 189
QY      64  CACTCTGACACCA 77
DB      188  CAGCCTGAACATCA 175

RESULT 13
US-10-154-884B-5294/C
; Sequence 5294, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
```

```
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5294
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(333)
; OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-5294
```

```
Query Match          31.7%; Score 27.6; DB 9; Length 333;
Best Local Similarity 60.8%; Pred. No. 3.1;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
```

```
QY 4 GAACCATCATTAATTGAAGTGAATTTTCTGGCCTGAGACTTGCAAGGAGGCAAGAAGA 63
Db 248 GAAAGTGCTGGGAGACACAACTTTTCTAGCTTCAGAAAGTCAGGAGCCCGCATCA 189
```

```
QY 64 CACTCTGACACCA 77
Db 188 CAGCCTGAACATCA 175
```

```
RESULT 14
US-09-833-381-916
; Sequence 916, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 916
; LENGTH: 587
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-916
```

```
Query Match          31.7%; Score 27.6; DB 9; Length 587;
Best Local Similarity 60.8%; Pred. No. 3.7;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
```

```
QY 4 GAACCATCATTAATTGAAGTGAATTTTCTGGCCTGAGACTTGCAAGGAGGCAAGAAGA 63
Db 132 GAAAGTGCTGGGAGACACAACTTTTCTAGCTTCAGAAAGTCAGGAGCCCGCATCA 191
```

```
QY 64 CACTCTGACACCA 77
Db 192 CAGCCTGAACATCA 205
```

```
RESULT 15
US-09-833-381-917
; Sequence 917, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 917
; LENGTH: 1003
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-917
```

```
Query Match          31.7%; Score 27.6; DB 9; Length 1003;
Best Local Similarity 60.8%; Pred. No. 4.4;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
```

```
QY 4 GAACCATCATTAATTGAAGTGAATTTTCTGGCCTGAGACTTGCAAGGAGGCAAGAAGA 63
Db 178 GAAAGTGCTGGGAGACACAACTTTTCTAGCTTCAGAAAGTCAGGAGCCCGCATCA 237
QY 64 CACTCTGACACCA 77
Db 238 CAGCCTGAACATCA 251
```

```
Search completed: March 13, 2004, 05:30:54
Job time : 31.53 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 12, 2004, 13:19:34 ; Search time 277.03 Seconds

(without alignments)
9378.080 Million cell updates/sec

Title: US-09-966-880A-11

Perfect score: 87

Sequence: 1 agagacacatcatatga.....ctggacacacatgacag 87

Scoring table: IDENTITY_NUC

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estov:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hrc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hrc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 87 | 100.0 | 535 | 14 | CD707143 EST23670 |
| 2 | 87 | 100.0 | 693 | 12 | BG757392 602711022 |
| 3 | 87 | 100.0 | 1201 | 13 | BX402063 BX402063 |
| 4 | 85.4 | 98.2 | 820 | 12 | BG757089 602715124 |

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 5 | 84 | 96.6 | 541 | 10 | BF238155 |
| 6 | 84 | 96.6 | 743 | 12 | BG686133 |
| 7 | 84 | 96.6 | 942 | 12 | BF975166 |
| 8 | 71 | 81.6 | 953 | 13 | BO065440 |
| 9 | 71 | 81.6 | 1052 | 13 | BO055935 |
| 10 | 70.4 | 80.9 | 872 | 12 | BG758510 |
| 11 | 39.2 | 45.1 | 1201 | 9 | AL559877 |
| 12 | 38.8 | 44.6 | 853 | 13 | BX464579 |
| 13 | 37 | 42.5 | 889 | 12 | BG686876 |
| 14 | 31.6 | 36.3 | 930 | 28 | CC198363 |
| 15 | 30.8 | 35.4 | 120 | 10 | BE071397 |
| 16 | 30 | 34.5 | 192 | 29 | CG530782 |
| 17 | 30 | 34.5 | 196 | 29 | CG656390 |
| 18 | 30 | 34.5 | 204 | 29 | CG545883 |
| 19 | 30 | 34.5 | 212 | 29 | CG578867 |
| 20 | 30 | 34.5 | 590 | 28 | BG620575 |
| 21 | 29.6 | 34.0 | 702 | 12 | BG757541 |
| 22 | 29.4 | 33.8 | 329 | 28 | AZ395639 |
| 23 | 29.4 | 33.8 | 508 | 28 | AZ158608 |
| 24 | 29.2 | 33.6 | 272 | 13 | BU732934 |
| 25 | 29 | 33.3 | 538 | 28 | AQ779004 |
| 26 | 29 | 33.3 | 770 | 29 | CG993610 |
| 27 | 28.8 | 33.1 | 612 | 29 | CE326757 |
| 28 | 28.8 | 33.1 | 622 | 29 | CE030990 |
| 29 | 28.8 | 33.1 | 661 | 14 | CB419607 |
| 30 | 28.8 | 33.1 | 704 | 14 | CA170646 |
| 31 | 28.8 | 33.1 | 736 | 28 | AZ519070 |
| 32 | 28.8 | 33.1 | 1002 | 13 | BQ709239 |
| 33 | 28.6 | 32.9 | 262 | 9 | AV388203 |
| 34 | 28.6 | 32.9 | 375 | 28 | AZ085135 |
| 35 | 28.6 | 32.9 | 896 | 14 | CB627885 |
| 36 | 28.4 | 32.6 | 473 | 28 | AQ859236 |
| 37 | 28.4 | 32.6 | 782 | 28 | BH284924 |
| 38 | 28.4 | 32.6 | 947 | 28 | CC537188 |
| 39 | 28.2 | 32.4 | 557 | 29 | AZ106106 |
| 40 | 28.2 | 32.4 | 628 | 28 | BX128009 |
| 41 | 28 | 32.2 | 442 | 28 | AQ276930 |
| 42 | 28 | 32.2 | 490 | 29 | CG722866 |
| 43 | 28 | 32.2 | 599 | 14 | CB243900 |
| 44 | 28 | 32.2 | 626 | 28 | CC156720 |
| 45 | 28 | 32.2 | 803 | 29 | CG093969 |

ALIGNMENTS

RESULT 1
CD707143 535 bp mRNA linear EST 25-JUN-2003
DEFINITION EST23670 human nasopharynx Homo sapiens cDNA, mRNA sequence.
CD707143
VERSION CD707143.1 GI:32237773
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G., and Zeng, Y.-X.
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL Unpublished (2003)
COMMENT Contact: Yixin Zeng
Cancer Center
Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China
Tel: 86-1380-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@zsusm.edu.cn.
Location/Qualifiers
1. 535
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/issue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"

ORIGIN

Query Match 100.0%; Score 87; DB 14; Length 535;
Best Local Similarity 100.0%; Pred. No. 4,8e-19;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGACCATCATTAATTAAGTGAAGATTTTCTGGCCTGAGACTTGACGAGGACGAGA 60
DB 42 AGAAGACCATCATTAATTAAGTGAAGATTTTCTGGCCTGAGACTTGACGAGGACGAGA 101
QY 61 AGACACTCTGGACACCACTATGACAG 87
DB 102 AGACACTCTGGACACCACTATGACAG 128

RESULT 2 693 bp mRNA linear EST 15-MAY-2001
BG757392
LOCUS 602711022F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4851580 5',
DEFINITION mRNA sequence.

ACCESSION BG757392.1 GI:14068045
VERSION BG757392.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS 1 (bases 1 to 693)
TITLE NIH-MGC http://mhc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM1694 row: k column: 05
High quality sequence stop: 693.
Location/Qualifiers

FEATURES

source

1..693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4851580"
/issue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GCGACGAG(G). State-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 87; DB 12; Length 693;
Best Local Similarity 100.0%; Pred. No. 5,2e-19;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGAACCATCATTAATTAAGTGAAGATTTTCTGGCCTGAGACTTGACGAGGACGAGA 60
|||||
|||||

DB 2 AAGAACCATCATTAATTAAGTGAAGATTTTCTGGCCTGAGACTTGACGAGGACGAGA 61
QY 61 AGACACTCTGGACACCACTATGACAG 87
DB 62 AGACACTCTGGACACCACTATGACAG 88

RESULT 3 1201 bp mRNA linear EST 13-MAY-2003
BX402063
LOCUS BX402063 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CSODL012YD18 5-PRIME, mRNA sequence.

ACCESSION BX402063
VERSION BX402063
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 1201)
TITLE Li W.B., Gruber C., Jesse J. and Polayes D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: sequef@genoscope.cns.fr. Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6672.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODL012DB09QPLcluster=6672.r. Contact :
Feng Liang Email : fliang@life tech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODL012DB09QPL.
Location/Qualifiers

FEATURES

source

1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODL012YD18"
/cell_type="B CELLS (RAMOS CELL LINE)"
/clone_lib="RAMOS CELL LINE"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 100.0%; Score 87; DB 13; Length 1201;
Best Local Similarity 100.0%; Pred. No. 6,2e-19;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGAACCATCATTAATTAAGTGAAGATTTTCTGGCCTGAGACTTGACGAGGACGAGA 60
DB 73 AAGAACCATCATTAATTAAGTGAAGATTTTCTGGCCTGAGACTTGACGAGGACGAGA 132
QY 61 AGACACTCTGGACACCACTATGACAG 87
DB 133 AGACACTCTGGACACCACTATGACAG 159

RESULT 4 820 bp mRNA linear EST 15-MAY-2001
BG757089
LOCUS BG757089
DEFINITION 602715124F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4855517 5',
mRNA sequence.
ACCESSION BG757089
VERSION BG757089.1 GI:14067742
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 820)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

cDNA library Preparation: Ling Hong/Rubin Laboratory

cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: LNCM1704 row: 0 column: 06

High quality sequence stop: 675.

Location/Qualifiers

1. 820

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:485517"

/lab_host="primary B-cells from tonsils (cell line)"

/note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by Ling

Hong in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

ORIGIN

Query Match 98.2%; Score 85.4; DB 12; Length 820;

Best Local Similarity 98.9%; Pred. No. 2e-18;

Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 60

DB 6 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 65

QY 61 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 87

DB 66 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 92

QY 61 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 87

DB 66 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 92

QY 61 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 87

DB 66 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 92

QY 61 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 87

DB 66 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 92

QY 61 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 87

DB 66 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 92

QY 61 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 87

DB 66 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 92

QY 61 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 87

DB 66 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 92

QY 61 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 87

DB 66 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 92

QY 61 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 87

DB 66 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 92

QY 61 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 87

DB 66 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 92

QY 61 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 87

DB 66 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 92

QY 61 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 87

DB 66 AGAAGACCATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 92

FEATURES

source

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>

Plate: LNCM695 row: p column: 20

High quality sequence stop: 541.

Location/Qualifiers

1. 541

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4054915"

/tissue_type="primary B-cells from tonsils (cell line)"

/lab_host="DH10B (phage-resistant)"

/note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by Ling

Hong in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

ORIGIN

Query Match 96.6%; Score 84; DB 10; Length 541;

Best Local Similarity 100.0%; Pred. No. 5.2e-18;

Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GAACATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 63

DB 2 GAACATCATTAATTAAGTGAATTTCTGCGCTGACACTTGACGAGGACCAAGA 61

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

QY 64 CACTTGACACCACTATGACAG 87

DB 62 CACTTGACACCACTATGACAG 85

FEATURES

source

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>

Plate: LNCM695 row: p column: 20

High quality sequence stop: 541.

Location/Qualifiers

1. 541

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4054915"

/tissue_type="primary B-cells from tonsils (cell line)"

/lab_host="DH10B (phage-resistant)"

/note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by Ling

Hong in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

Location/Qualifiers

1. 743

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4766234"

/tissue_type="primary B-cells from tonsils (cell line)"

/lab_host="DH10B (phage-resistant)"

Note: this is a NIH_MGC Library."

Location/Qualifiers

1. 743

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4766234"

/tissue_type="primary B-cells from tonsils (cell line)"

/lab_host="DH10B (phage-resistant)"

Note: this is a NIH_MGC Library."

Location/Qualifiers

1. 743

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4766234"

/tissue_type="primary B-cells from tonsils (cell line)"

/lab_host="DH10B (phage-resistant)"

Note: this is a NIH_MGC Library."

Location/Qualifiers

1. 743

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4766234"

/tissue_type="primary B-cells from tonsils (cell line)"

/lab_host="DH10B (phage-resistant)"

Note: this is a NIH_MGC Library."

Location/Qualifiers

/clone.lib="NIH_MGC_48"
 /note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."

ORIGIN

Query Match 96.6%; Score 84; DB 12; Length 743;

Best Local Similarity 100.0%; Pred. No. 5.7e-18;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GAACCATCATTTAATGAAGTGAATTTTCTGGCTGAGACTTGCAGGAGGCAAGA 63
 DB 2 GAACCATCATTTAATGAAGTGAATTTTCTGGCTGAGACTTGCAGGAGGCAAGA 61
 QY 64 CACTCTGACACCACTATGACAG 87
 DB 62 CACTCTGACACCACTATGACAG 85

RESULT 7

BP975166 942 bp mRNA linear EST 22-JAN-2001

LOCUS 60224465791 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4335639 5',

DEFINITION mRNA sequence.

ACCESSION BP975166

VERSION BP975166.1 GI:12342381

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgc.nci.nih.gov/.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cga@bbs-remail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LNCM1207 row: a column: 16

High quality sequence stop: 707.

Location/Qualifiers

1. 942

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4335639"

/lab_host="PH10B (phage-resistant)"

/note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."

ORIGIN

Query Match

96.6%; Score 84; DB 10; Length 942;

Best Local Similarity 100.0%; Pred. No. 6.2e-18;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GAACCATCATTTAATGAAGTGAATTTTCTGGCTGAGACTTGCAGGAGGCAAGA 63
 DB 2 GAACCATCATTTAATGAAGTGAATTTTCTGGCTGAGACTTGCAGGAGGCAAGA 61
 QY 64 CACTCTGACACCACTATGACAG 87
 DB 62 CACTCTGACACCACTATGACAG 85

RESULT 8

B0065440 953 bp mRNA linear EST 02-APR-2002

LOCUS AGENCOURT 6855061 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5929977

DEFINITION 5', mRNA sequence.

ACCESSION B0065440

VERSION B0065440.1 GI:19894486

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 953)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cga@bbs-remail.nih.gov

Tissue Procurement: Lou Staudt

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LNCM2108 row: p column: 10

High quality sequence stop: 634.

Location/Qualifiers

1. 953

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5929977"

/lab_host="PH10B (phage-resistant)"

/note="Organ: lymph; Vector: pOTB7; Site: 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 library."

ORIGIN

Query Match 81.6%; Score 71; DB 13; Length 953;

Best Local Similarity 100.0%; Pred. No. 1.8e-13;
 Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 TTGAAGTGAAGATTTTCTGGCTGAGACTTGCAGGAGGCAAGAAGACACTCTGACACC 76
 DB 1 TTGAAGTGAAGATTTTCTGGCTGAGACTTGCAGGAGGCAAGAAGACACTCTGACACC 60
 QY 77 ACTATGACAG 87
 DB 61 ACTATGACAG 71

RESULT 9

B0055935 1052 bp mRNA linear EST 29-MAR-2002
 LOCUS B0055935
 DEFINITION AGENCOURT 6796291 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5808181
 5', mRNA sequence.
 ACCESSION B0055935
 VERSION B0055935.1 GI:19815262
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1052)
 NIH-MGC <http://mge.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgsapbs-remail.nih.gov
 Tissue Procurement: Lou Staudt
 cDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L10CM2051 row: m column: 14
 High quality sequence stop: 665.
 Location/Qualifiers
 1..1052
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5808181"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."

ORIGIN
 Query Match 81.6%; Score 71; DB 13; length 1052;
 Best Local Similarity 100.0%; Pred. No. 1.8e-13;
 Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 TTGAAGTGAATTTTTCGGCCTGAGACTTCAGAGAGGCAAGACACTCTGGAACC 76
 Db 1 TTGAAGTGAATTTTTCGGCCTGAGACTTCAGAGAGGCAAGACACTCTGGAACC 60
 QY 77 ACTATGACAG 87
 Db 61 ACTATGACAG 71

RESULT 10 872 bp mRNA linear EST 15-MAY-2001
 LOCUS B0758510
 DEFINITION 60271272121 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853069 5',
 mRNA sequence.
 ACCESSION B0758510
 VERSION B0758510.1 GI:14069163
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 872)
 NIH-MGC <http://mge.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgsapbs-remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L10CM1638 row: i column: 06
 High quality sequence stop: 836.
 Location/Qualifiers
 1..872
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4853069"
 /tissue_type="primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 80.9%; Score 70.4; DB 12; length 872;
 Best Local Similarity 98.6%; Pred. No. 2.8e-13;
 Matches 71; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 16 ATTGAAGTGAATTTTTCGGCCTGAGACTTCAGAGAGGCAAGACACTCTGGAACC 75
 Db 2 ATTGAAGTGAATTTTTCGGCCTGAGACTTCAGAGAGGCAAGACACTCTGGAACC 61
 QY 76 CACTATGACAG 87
 Db 62 CACTATGACAG 73

RESULT 11 1201 bp mRNA linear EST 31-MAY-2003
 LOCUS AL559877
 DEFINITION AL559877 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
 clone GS0D003YB14 5-PRIME, mRNA sequence.
 ACCESSION AL559877
 VERSION AL559877.2 GI:31284008
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1201)
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:12905793.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: sequef@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 6672.r For
 more information about this cluster, see
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=GS0D003YB14)
 cgi-bin/cluster.cgi?seq=GS0D003YB14
 Feng Liang Email: fliang@lifeotech.com URL:

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D0003DA07Qp1.

FEATURES
Source

1..1201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D0003YB14"
/issue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT 6, 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match 45.1%; Score 39.2; DB 9; Length 1201;
Best Local Similarity 73.5%; Pred. No. 0.015; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 18;

QY 20 AAGTGAATTTTCTGCGCTGAGACTTGACAGAGGAGAGACACTTGACACCACT 79
DB 32 AAGCAGCTGTGACCGGTCCGGAATTCCTCGGATGGAGAGACACTTGACACCACT 91
QY 80 ATGCACAG 87
DB 92 ATGCACAG 99

RESULT 12

EX464579 853 bp mRNA linear EST 22-MAY-2003
LOCUS EX464579 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
DEFINITION clone CS0D0003YB14 5-PRIME, mRNA sequence.

ACCESSION BX464579
VERSION BX464579.1 GI:31031641
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Fukuyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 853)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: sequef@genoscope.cns.fr; Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6672.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS1D0001ZF10Qp1&cluster=6672.r. Contact :
Feng liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS1D0001ZF10Qp1.

FEATURES

Source

1..853
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D0003YB14"
/issue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT 6, 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match 44.6%; Score 38.8; DB 13; Length 853;
Best Local Similarity 72.1%; Pred. No. 0.019;
Matches 49; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

QY 20 AAGTGAATTTTCTGCGCTGAGACTTGACAGAGGAGAGACACTTGACACCACT 79
DB 27 AAGCAGCTGTGACCGGTCCGGAATTCCTCGGATGGAGAGACACTTGACACCACT 86
QY 80 ATGCACAG 87
DB 87 ATGCACAG 94

RESULT 13

BG686876 889 bp mRNA linear EST 01-MAY-2001
LOCUS 602650861F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4763247 5',
DEFINITION mRNA sequence.

ACCESSION BG686876
VERSION BG686876.1 GI:13918273
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cga@bbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA library preparation: Ling Hong/Rubin Laboratory
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM1618 row: 5 column: 16
High quality sequence start: 6
High quality sequence stop: 727.
Note: this is a NIH_MGC Library."

FEATURES

Source

1..889
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4763247"
/issue_type="Primary B-cells from tonsils (cell line)"
/lab_host="PH1B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7, Site:1; XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald W. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 42.5%; Score 37; DB 12; Length 889;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GGAGGAGAGAGACACTTGACACCACTATGACAG 87
DB 9 GGAGGAGAGAGACACTTGACACCACTATGACAG 45

RESULT 14

```

sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
20202663
PUBMED
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?it=st2-RC-BT0506-240
400-016-905&t3=2000-04-24&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 26
High quality sequence stop: 250.
FEATURES
Source
Location/Qualifiers
1..250
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BT0506"
/note="Organ: breast; Vector: puc18, Site_1: SmaI, Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
ORIGIN
Query Match 35.4%; Score 30.8; DB 10; Length 250;
Best Local Similarity 63.5%; Pred.No.7.2;
Matches 47; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 4 GAACATCATTAATTGAGTGAATTTTTCGGCCTGAGACTTGCAGGAGGCAAGAAGA 63
Db 123 GAAGGGTGTTGGAGAGGGGAGCACTTTTCTAGCTTCAGAAAGTCAGGAGCCAGATCA 64
QY 64 CACTTGGACACCA 77
Db 63 CAGCCTGAACATCA 50
Search completed: March 13, 2004, 00:38:59
Job time : 280.155 secs

```


DB 1 AGAGACCATCATTAATTGAAAGTGAAGATTTTCTGGCCTGAGACTTGAGGAGGCAAGA 60
QY 61 AGACACTCTGGACACCACTATGACAG 87
DB 61 AGACACTCTGGACACCACTATGACAG 87

RESULT 2
BD016833 2818 bp DNA linear PAT 27-AUG-2002
LOCUS Novel cytidine deaminase.
DEFINITION BD016833
ACCESSION BD016833.1 GI:22558009
VERSION JP 2001245669-A/6.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Honjo, T. and Muramatsu, M.
TITLE Novel cytidine deaminase
JOURNAL Patent: JP 2001245669-A 6 11-SEP-2001;
JAPAN TOBACCO INC, TASUKU HONJO
OS Homo sapiens (human)
PN JP 2001245669-A/6
PD 11-SEP-2001
PF 28-MAR-2000 JP 2000092981
PI TASUKU HONJO, MASAMICHI MURAMATSU
PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12,
A61P17/00,
PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12N9/78, C12P21/02, C12P21/08, C12N1/21, C12R1/19, PC
(C12N5/10, C12R1/91), C12N15/00, C12N5/00, (C12N5/00, C12R1/91) CC
FH Key Location/Qualifiers
FT 5'UTR (1), (79)
FT CDS (80), (676)
FT 3'UTR (677), (2818).
FEATURES
source Location/Qualifiers
1..2818
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 87; DB 6; Length 2818;
Best Local Similarity 100.0%; Pred. No. 8,7e-41;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGACCATCATTAATTGAAAGTGAAGATTTTCTGGCCTGAGACTTGAGGAGGCAAGA 60
DB 1 AGAGACCATCATTAATTGAAAGTGAAGATTTTCTGGCCTGAGACTTGAGGAGGCAAGA 60

QY 61 AGACACTCTGGACACCACTATGACAG 87
DB 61 AGACACTCTGGACACCACTATGACAG 87

RESULT 3
BD016834 5514 bp DNA linear PAT 27-AUG-2002
LOCUS Novel cytidine deaminase.
DEFINITION BD016834
ACCESSION BD016834.1 GI:22558010
VERSION JP 2001245669-A/7.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Honjo, T. and Muramatsu, M.
TITLE Novel cytidine deaminase
JOURNAL Patent: JP 2001245669-A 7 11-SEP-2001;

COMMENT JAPAN TOBACCO INC, TASUKU HONJO
OS Homo sapiens (human)
PN JP 2001245669-A/7
PD 11-SEP-2001
PF 28-MAR-2000 JP 2000092981
PI TASUKU HONJO, MASAMICHI MURAMATSU
PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12,
A61P17/00,
PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12N9/78, C12P21/02, C12P21/08, C12N1/21, C12R1/19, PC
(C12N5/10, C12R1/91), C12N15/00, C12N5/00, (C12N5/00, C12R1/91) CC
FH Key Location/Qualifiers
FT 5'UTR (1), (1031)
FT exon (1032), (1118)
FT intron (1119), (5514).
FEATURES
source Location/Qualifiers
1..5514
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 87; DB 6; Length 5514;
Best Local Similarity 100.0%; Pred. No. 8e-41;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGACCATCATTAATTGAAAGTGAAGATTTTCTGGCCTGAGACTTGAGGAGGCAAGA 60
DB 1032 AGAGACCATCATTAATTGAAAGTGAAGATTTTCTGGCCTGAGACTTGAGGAGGCAAGA 1091

QY 61 AGACACTCTGGACACCACTATGACAG 87
DB 1092 AGACACTCTGGACACCACTATGACAG 1118

RESULT 4
BD016860 11204 bp DNA linear PAT 27-AUG-2002
LOCUS Novel cytidine deaminase.
DEFINITION BD016860
ACCESSION BD016860.1 GI:22558036
VERSION JP 2001245669-A/33.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Honjo, T. and Muramatsu, M.
TITLE Novel cytidine deaminase
JOURNAL Patent: JP 2001245669-A 33 11-SEP-2001;
JAPAN TOBACCO INC, TASUKU HONJO
OS Homo sapiens (human)
PN JP 2001245669-A/33
PD 11-SEP-2001
PF 28-MAR-2000 JP 2000092981
PI TASUKU HONJO, MASAMICHI MURAMATSU
PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12,
A61P17/00,
PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12N9/78, C12P21/02, C12P21/08, C12N1/21, C12R1/19, PC
(C12N5/10, C12R1/91), C12N15/00, C12N5/00, (C12N5/00, C12R1/91) CC
FH Key Location/Qualifiers
1..11204
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 87; DB 6; Length 11204;
Best Local Similarity 100.0%; Pred. No. 7.4e-41;

Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAACCATCATTAATTGAGATTTTCTGGCTGAGACTTGACGAGGAGCAAGA 60
 DB 442 AGAGAACCATCATTAATTGAGATTTTCTGGCTGAGACTTGACGAGGAGCAAGA 501

QY 61 AGACACTCTGGACACCACTATGACAG 87
 DB 502 AGACACTCTGGACACCACTATGACAG 528

RESULT 5
 AB040430 11204 bp DNA linear PRI 03-OCT-2000
 LOCUS AB040430
 DEFINITION Homo sapiens AID gene for activation-induced cytidine deaminase, complete cds.

ACCESSION AB040430
 VERSION AB040430.1 GI:99888407
 KEYWORDS AID; activation-induced cytidine deaminase.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Muto, T., Muramatsu, M., Taniguchi, M., Kinoshita, K. and Honjo, T.
 TITLE Isolation, tissue distribution, and chromosomal localization of the human activation-induced cytidine deaminase (AID) gene
 JOURNAL Genomics 68 (1), 85-88 (2000)
 MEDLINE 20408890
 PUBMED 10950930
 REFERENCE
 AUTHORS Revy, P., Muto, T., Levy, Y., Gelsmann, F., Plehant, A., Sanal, O., Catelan, N., Forveille, M., Dufourcq-Iagelouse, R., Gennery, A., Tezcan, I., Ercoy, F., Kayseril, H., Ugazio, A.G., Brousse, N., Muramatsu, M., Notarangelo, L.D., Kinoshita, K., Honjo, T., Fischer, A. and Durandy, A.
 TITLE Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2)
 JOURNAL Cell 102 (5), 565-575 (2000)
 MEDLINE 20460541
 PUBMED 11007475

FEATURES
 source
 1. (bases 1 to 11204)
 Muto, T., Muramatsu, M., Taniguchi, M., Kinoshita, K. and Honjo, T.
 Direct Submission
 Submitted (18-MAR-2000) Tasuku Honjo, Kyoto University, Department of Medical Chemistry, Faculty of Medicine, Yoshida, Sakyo-ku, Kyoto, Kyoto 606-8501, Japan (E-mail: honjo@med.kyoto-u.ac.jp, Tel:81-75-753-4371 (ex.4371), Fax:81-75-753-4388)
 Location/Qualifiers
 1. 11204
 /organism="Homo sapiens"
 /mol_type="Genomic DNA"
 /db_xref="taxon:9606"
 join(521..528,6280..6427,7807..8077,8371..8486,8956..9009)
 /gene="AID"
 join(521..528,6280..6427,7807..8077,8371..8486,8956..9009)
 /gene="AID"
 /codon_start=1
 /product="activation-induced cytidine deaminase"
 /protein_id="BAB12720.1"
 /db_xref="GI:99888408"
 /translation="MDSLNNRRKFLYQFNKVRKGRRETYLYCYVVRKDSATSFSLDFGYLRNKGCHVEFLFRLYISMDLDPGRCRYVTSWSPCDCAHVADELRGNP NLSIRIFARLYFCEDRKAEPEGRLRLRAGVQIAIMFKYFYFCMTFVNNHRTK AMEGHNSVLSRLRRLILPLIVVDLRARFTLIG"

ORIGIN
 Query Match 100.0%; Score 87; DB 9; Length 11204;
 Best Local Similarity 100.0%; Pred. No. 7,4e-41;
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAACCATCATTAATTGAGATTTTCTGGCTGAGACTTGACGAGGAGCAAGA 60
 DB 442 AGAGAACCATCATTAATTGAGATTTTCTGGCTGAGACTTGACGAGGAGCAAGA 501

DB 442 AGAGAACCATCATTAATTGAGATTTTCTGGCTGAGACTTGACGAGGAGCAAGA 501

QY 61 AGACACTCTGGACACCACTATGACAG 87
 DB 502 AGACACTCTGGACACCACTATGACAG 528

RESULT 6
 AC092184 71132 bp DNA linear PRI 12-JUN-2002
 LOCUS AC092184
 DEFINITION Homo sapiens 12 BAC RP11-438L7 (Roswell Park Cancer Institute Human BAC library) complete sequence.

ACCESSION AC092184
 VERSION AC092184.7 GI:21206067
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Al-Osman, F.R., Allen, C., Alsbrooks, S.L., Amaral, H.C., Are, J.R., Ayala, M., Banks, T., Barbieri, J., Benton, J., Bimaye, K., Blankensburg, K., Bonnin, D., Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burck, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Dublin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, K., Harris, K., Hart, M., Havlak, F., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hulik, S., Hume, J., Ioshikhes, I., Jackson, J.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kurehli, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Loussed, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maneswari, M., Mapa, P., Marondei, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Matzner, E., Meador, M.P., Meador, M., Mei, G., Merscher, S., Metzner, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, E., Nwokenkwo, S., Ogih, M., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okunou, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Peters, L., Pickens, R., Primus, B., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shooharati, N., Sisson, I., Sodergren, E., Sonaike, I., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabori, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansy, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Uemami, K., Vasquez, L., Vera, V., Vallation, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Kuchelapatti, R., Weinstein, G., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 71132)
 Worley, K.C.
 Direct Submission
 Submitted (25-JUN-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 71132)
 Worley, K.C.
 Direct Submission

JOURNAL Submitted (18-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS Morley, K.C.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS Morley, K.C.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On May 25, 2002 this sequence version replaced gi:20901754. INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

FEATURES
source Location/Qualifiers

1..71132
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-438L7"
complement(1..1448)
/note="overlaps bases 1..1448 of clone AC092490"
/function="Clone overlap"
187..408
/standard_name="57233"
439..560
/standard_name="92005"
774..881
/rpt_family="MIR"
903..1190
/rpt_family="Alusg"
repeat_region
1191..1213
repeat_region

STS /rpt_family="AT_rich"
1591..1807
/standard_name="6612"
STS 1744..1819
/standard_name="6198"
repeat_region 1966..2224
/rpt_family="Alusg"
repeat_region 3296..3328
/rpt_family="(TTTC)n"
repeat_region 4097..4249
/rpt_family="Alusg"
repeat_region 4360..4654
/rpt_family="AluY"
repeat_region complement(5140..5262)
/rpt_family="FLAM_C"
repeat_region 5275..5304
/rpt_family="AT_rich"
repeat_region 5305..5563
/rpt_family="AluO"
repeat_region 5564..5599
/rpt_family="AT_rich"
repeat_region 5644..5733
/rpt_family="MSTD"
repeat_region 5734..6012
/rpt_family="Alusg"
repeat_region 6020..6336
/rpt_family="AlusX"
repeat_region 6337..6629
/rpt_family="MSTD"
repeat_region 6630..6650
/rpt_family="(CNA)n"
repeat_region complement(6711..6817)
/rpt_family="XIR"
repeat_region complement(7033..7327)
/rpt_family="AlusX"
repeat_region 7607..7745
/rpt_family="FLAM_A"
repeat_region 7980..8047
/rpt_family="MIR"
repeat_region complement(8050..8549)
/rpt_family="LTR47A"
repeat_region 8567..8698
/rpt_family="MIR"
repeat_region complement(8815..9101)
/rpt_family="Alusg"
repeat_region complement(11227..12175)
/rpt_family="LTR5"
STS 12625..12900
/standard_name="87432"
repeat_region complement(13547..13664)
/rpt_family="LIMB5"
repeat_region complement(13665..13979)
/rpt_family="AluY"
repeat_region complement(13980..14193)
/rpt_family="LIMB5"
repeat_region 14622..14924
/rpt_family="AluY"
repeat_region complement(15813..17136)
/rpt_family="LIPAB8"
repeat_region complement(17272..17395)
/rpt_family="FLAM_C"
repeat_region 17426..17575
/rpt_family="MIR"

Query Match 100.0%; Score 87; DB 9; Length 71132;
Best Local Similarity 100.0%; Pred. No. 5,9e-41;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGAGAACCATCATTAATGAAGAGATTTTCTGCGCCGAGACTTGACGAGGAGGAGAGA 60
DB 34990 AGAGAACCATCATTAATGAAGAGATTTTCTGCGCCGAGACTTGACGAGGAGGAGAGA 35049
OY 61 AGACACTCTGACACACATATGACAG 87

DB 35050 AGACACTGTGACACCACTATGACAG 35076

RESULT 7 1828 bp mRNA linear PRI 03-OCT-2003
LOCUS BC006296

DEFINITION Homo sapiens activation-induced cytidine deaminase, mRNA (CDNA
clone MGC:12911 IMAGE:4054915), complete cds.

ACCESSION BC006296
VERSION BC006296.2 GI:33871601

KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM

REFERENCE 1 (bases 1 to 1828)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shemmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marinina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalton, D.E.,
Scherer, A., Schein, J.E., Jones, S.J., and Marra, M.A.

TITLE human and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 1828)
Strausberg, R.
Direct Submission
Submitted (09-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:13623400.

COMMENT Contact: MGC help desk
Email: cgapsb-r@mail.nih.gov
Tissue Procurement: Louis Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgrl.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Binkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lari, P., Legaspi, R.,
Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantipop, S., Thomas, P.J., Touchman, J.W.,
Tsurganov, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 17 Row: a Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 10190699.

FEATURES

SOURCE

Location/Qualifiers
1..1828
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:12911 IMAGE:4054915"
/issue_type="Primary B-cells from Tonsils"
/clone_id="NIH_MGC_48"
/lab_host="DH10B-R"
/note="Vector: pOT87"
1..1828
/gene="AIDCA"
/note="Synonyms: AID, HIGM2, CDA2, ARP2"
/db_xref="LocusID:57379"
/db_xref="MIM:605257"
77..673
/codon_start=1
/product="activation-induced cytidine deaminase"
/protein_id="AAH06296.1"
/db_xref="GI:13623401"
/db_xref="LocusID:57379"
/translation="MDSILNRKRPXYOFANVAKGRRETYCYVKRDSATPSL
DFGVLNRKNGCHIELFLRYSIDMDIDRCYVMTPTSKSPCYDCAHYADFLRGP
NLSRLFTALVIFCEDKAPBGLRLHRAGVQIATMTEDYFCMTVEVNERIFK
AMEGLHENSVALSRQRLRIILPLVEVDLDRAFTVLGL"

CDS

ORIGIN

Query Match 96.6%; Score 84; DB 9; Length 1828;
Best Local Similarity 100.0%; Pred. No. 5.7e-39; Indels 0; Gaps 0;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4 GAACCATCAATTAATTGAGTGAATTTCTGGCCCTGAGACTTGACAGGAGGAGAGA 63
1 GAACCATCAATTAATTGAGTGAATTTCTGGCCCTGAGACTTGACAGGAGGAGAGA 60
64 CACTCTGACACCACTATGACAG 87
61 CACTCTGACACCACTATGACAG 84

RESULT 8 2791 bp mRNA linear PRI 03-OCT-2000
LOCUS AB040431
DEFINITION Homo sapiens AID mRNA for activation-induced cytidine deaminase,
complete cds.
AB040431.1 GI:9988409
ACCESSION AB040431
VERSION AID; activation-induced cytidine deaminase; Human AID.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (sites)
Muro, T., Muramatsu, M., Taniwaki, M., Kinoshita, K. and Honjo, T.
Isolation, tissue distribution, and chromosomal localization of the
human activation-induced cytidine deaminase (AID) gene
Genomics 68 (1), 85-88 (2000)
JOURNAL MEDLINE 20408690
PUBMED 10950930
REFERENCE 2 (sites)
Remy, P., Muro, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O.,
Catalan, N., Forveille, M., Dufourcq-Lageat, R., Gennery, A.,
Tescan, I., Eysy, F., Kayserili, H., Ugazio, A.G., Brousse, N.,
Muramatsu, M., Notarangelo, L.D., Kinoshita, K., Honjo, T., Fischer, A.
and Durandy, A.
Activation-induced cytidine deaminase (AID) deficiency causes the
autosomal recessive form of the Hyper-IGM syndrome (HIGM2)
Cell 102 (5), 565-575 (2000)
JOURNAL MEDLINE 20460541
PUBMED 11007475
REFERENCE 3 (bases 1 to 2791)
Muro, T., Muramatsu, M., Taniwaki, M., Kinoshita, K. and Honjo, T.
Direct Submission


```
/note="L1MA3 repeat: matches 5568. .6304 of consensus"
repeat_region
2073. .2093
/note="L1MA3 repeat: matches 5257. .5277 of consensus"
repeat_region
2279. .2344
/note="33 copies 2 mer aa 69% conserved"
repeat_region
2560. .2621
/note="31 copies 2 mer at 72% conserved"
misc_feature
3248. .3483
/note="match: GSS: Em:AQ059990"
repeat_region
3442. .3571
/note="MIR repeat: matches 89. .249 of consensus"
misc_feature
3490. .3608
/note="match: GSS: Em:AQ059990"
repeat_region
5092. .5763
/note="L1PA9 repeat: matches 5178. .5854 of consensus"
repeat_region
5779. .6067
/note="L1PA9 repeat: matches 5859. .6163 of consensus"
repeat_region
7865. .7949
/note="L1T1H repeat: matches 72. .156 of consensus"
repeat_region
8143. .8285
/note="L1T1H repeat: matches 404. .547 of consensus"
repeat_region
8387. .8806
/note="MSTB repeat: matches 1. .426 of consensus"
misc_feature
9159. .9604
/note="match: GSS: Em:AQ059871"
repeat_region
9562. .10457
/note="L1MB4 repeat: matches 5290. .6182 of consensus"
repeat_region
11290. .11394
/note="MIR repeat: matches 65. .179 of consensus"
repeat_region
11848. .12266
/note="MSTB repeat: matches 3. .426 of consensus"
misc_feature
14021. .14444
/note="match: GSS: Em:AQ19307"
repeat_region
14021. .14444
/note="MSTB repeat: matches 1. .426 of consensus"
repeat_region
15546. .15724
/note="MSTB repeat: matches 1. .426 of consensus"
repeat_region
15842. .15936
/note="MSTB repeat: matches 1. .426 of consensus"
repeat_region
16125. .18230
/note="L1PA1.5 repeat: matches 4040. .6151 of consensus"
misc_feature
18459. .18906
/note="match: GSS: Em:AQ240510"
misc_feature
18465. .18819
/note="match: GSS: Em:B31098"
misc_feature
18481. .18817
/note="match: GSS: Em:B31099"
complement(18659. .18945)
/note="match: GSS: Em:AQ527813"
18856. .19135
/note="match: GSS: Em:AQ527813"
repeat_region
19273. .19348
/note="L1U10 repeat: matches 5. .282 of consensus"
repeat_region
19357. .19418
/note="MERSB repeat: matches 2. .78 of consensus"
repeat_region
19402. .19767
/note="MERSB repeat: matches 494. .546 of consensus"
repeat_region
19402. .19767
/note="MERSB repeat: matches 71. .436 of consensus"
repeat_region
19768. .21133
/note="L1PA8 repeat: matches 2430. .3809 of consensus"
repeat_region
21125. .23453
/note="L1PA8 repeat: matches 3824. .6163 of consensus"
repeat_region
23488. .23568
/note="MERSB repeat: matches 5. .88 of consensus"
repeat_region
23852. .23909
/note="MERSB repeat: matches 1. .58 of consensus"
repeat_region
26262. .26380
/note="L1U10 repeat: matches 1. .311 of consensus"
repeat_region
28100. .28165
/note="MERSB repeat: matches 457. .553 of consensus"
```

```
repeat_region
28333. .28656
/note="L1USX repeat: matches 1. .312 of consensus"
repeat_region
28675. .28908
/note="MERS3 repeat: matches 148. .177 of consensus"
repeat_region
28909. .29004
/note="L1R29 repeat: matches 4. .100 of consensus"
repeat_region
29005. .29015
/note="MERS3 repeat: matches 21. .148 of consensus"
repeat_region
29817. .30046
/note="L1TIGER1 repeat: matches 2175. .2413 of consensus"
repeat_region
30047. .30342
/note="L1USC repeat: matches 1. .294 of consensus"
repeat_region
30343. .30860
/note="L1TIGER1 repeat: matches 1695. .2175 of consensus"
repeat_region
30855. .30900
/note="L1TIGER1 repeat: matches 198. .250 of consensus"
repeat_region
30985. .31132
/note="L1TIGER1 repeat: matches 23. .166 of consensus"
gene
32392. .33680
/gene="ba278020.1"
/pseudo
32392. .33680
/gene="ba278020.1"
/note="ba278020.1 (RBBP4 (retinoblastoma-binding protein
4) pseudogene)
match: CDNA: Em:AF073787 Em:AF090306 Em:X71810 Em:U35141
Em:U35142 Em:U35143 Em:X74262 Em:A19284 Em:AF097750
match: ESTs: Em:AA554337 Em:AA556655 Em:A1914754
Em:AA040411 Em:AA403752 Em:A114668 Em:AA12520
Em:AA402945 Em:AA354104 Em:AA313613 Em:AA62873
Em:AA705035 Em:AA263144 Em:AA337765 Em:AA506805
Em:AA007448 Em:AA429422 Em:AA362878 Em:R67309 Em:AA853720
Em:AA332449 Em:A1099022 Em:AA359717 Em:AA359673
Em:AA452848 Em:HS1237 Em:AA227133 Em:AA310336 Em:AA40303
Em:AA330258 Em:AA331743 Em:HA6555 Em:AA386086 Em:AA007492
Em:AA331529 Em:AL080041 Em:AA082962 Em:AA292584
Em:AA386066 Em:R39087 Em:AA281307 Em:AA382542 Em:AA13395
Em:AA730527 Em:AA355615 Em:AA171394 Em:AA67949
Em:AA615284 Em:A1138310 Em:A1060994 Em:N92003 Em:AA403348
Em:DS466 Em:AL038896
match: proteins: Sw:P90917 Sw:Q22468 Sw:Q22469 Sw:P90916
Tr:O94244 Sw:P39984 Sw:O14021 Sw:P43254 Tr:O43974
Sw:Q09028 Sw:Q24572 Sw:Q60972 Sw:Q60973 Tr:Q85U78
Sw:Q22466 Sw:Q22467 Tr:O00809 Sw:P34471 Sw:Q16576
Tr:O9Y825 Sw:Q22607"
/pseudo
/codon_start=1
/evidence=not_experimental
repeat_region
35461. .35982
/note="MERS1B repeat: matches 3. .617 of consensus"
repeat_region
36651. .37134
/note="L12 repeat: matches 2025. .2595 of consensus"
repeat_region
37135. .37451
/note="L1USX repeat: matches 1. .312 of consensus"
repeat_region
37452. .37603
/note="L12 repeat: matches 2594. .2745 of consensus"
repeat_region
37683. .37811
/note="MIR repeat: matches 68. .200 of consensus"
repeat_region
37812. .37845
/note="L17 copies 2 mer ca 88% conserved"
repeat_region
37889. .38025
/note="L1R45 repeat: matches 349. .486 of consensus"
repeat_region
40149. .40908
/note="L1MB7 repeat: matches 5406. .6173 of consensus"
repeat_region
40950. .41131
```

Query Match 24.1% Score 21; DB 9; Length 140120;
Best Local Similarity 100.0%; Pred No. 0.18;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 GGAGGCAAGAGCACTCTGG 71
Db 71571 GGAGGCAAGAGCACTCTGG 71591

| RESULT 11 | AC026139 | 142126 bp | DNA | linear | HTG 06-MAY-2001 |
|------------|---|-------------------------------|-----|--------|-----------------|
| LOCUS | AC026139 | | | | |
| DEFINITION | Homo sapiens chromosome 6 clone RP11-278U20 map 6, WORKING DRAFT | | | | |
| ACCESSION | AC026139 | SEQUENCE, 23 uncloned pieces. | | | |
| VERSION | AC026139.2 | GI:9994157 | | | |
| KEYWORDS | HTG, HTGS, PHASE1, HTGS_DRAFT. | | | | |
| SOURCE | Homo sapiens (human) | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | | | |
| TITLE | Birren, B., Linton, L., Nusbaum, C. and Lander, E. | | | | |
| JOURNAL | 1 (bases 1 to 142126) | | | | |
| REFERENCE | Homo sapiens chromosome 6, clone RP11-278U20 | | | | |
| AUTHORS | Unpublished | | | | |
| | 2 (bases 1 to 142126) | | | | |
| | Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F., Boguski, M., Boulanger, B., Brown, A., Burkett, G., Campyano, A., Castle, A., Chappel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearfield, K., Dewar, K., Diaz, J. S., Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardy, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karakas, A., Klein, J., Lacombe, K., Lamazares, R., Landers, T., Lehotzky, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Margulis, N., McCarthy, M., McGowan, P., McGuirk, A., McKernan, K., McPherson, R., Melnick, J., Meneus, L., Mihova, T., Miranda, C., Mlenka, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., Plesniak, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange, S., Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testa, S., Theodore, J., Tirrell, A., Travers, M., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, M. J., Young, G., Zairoun, J., Zimmer, A. and Zody, M. | | | | |
| | Direct Submission | | | | |
| | Submitted (19-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA | | | | |
| | On Sep 8, 2000 this sequence version replaced gi:7264209. | | | | |
| | All repeats were identified using RepeatMasker: | | | | |
| | Smit, A. F. A. & Green, P. (1996-1997) | | | | |
| | http://ftp.genome.washington.edu/RM/RepeatMasker.html | | | | |
| | Genome Center | | | | |
| | Center: Whitehead Institute/ MIT Center for Genome Research | | | | |
| | Center code: MIBR | | | | |
| | Web site: http://www-seq.wi.mit.edu | | | | |
| | Contact: sequence submissions@genome.wi.mit.edu | | | | |
| | Project Information | | | | |
| | Center project name: L8025 | | | | |
| | Center clone name: 278_U20 | | | | |
| | Summary Statistics | | | | |
| | Sequencing vector: M13; M7815; 100% of reads | | | | |
| | Chemistry: Dye-terminator Big Dye; 100% of reads | | | | |
| | Assembly program: Phrap; version 0.960721 | | | | |
| | Consensus quality: 132425 bases at least Q40 | | | | |
| | Consensus quality: 137265 bases at least Q30 | | | | |
| | Consensus quality: 139029 bases at least Q20 | | | | |
| | Insert size: 140000; agarose-efp | | | | |
| | Insert size: 139926; sum-of-contigs | | | | |
| | Quality coverage: 4.5 in Q20 bases; agarose-efp | | | | |
| | Quality coverage: 4.5 in Q20 bases; sum-of-contigs | | | | |
| | NOTE: This is a 'working draft' sequence. It currently consists of 23 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence | | | | |

```

* as soon as it is available and the accession number will
* be preserved.
1
16449: contig of 16449 bp in length
16450
16549: gap of 100 bp
16550
17637: contig of 1288 bp in length
17638
17937: gap of 100 bp
17938
19143: contig of 1206 bp in length
19144
19743: gap of 100 bp
19744
20877: contig of 1544 bp in length
20878
20888
23321: contig of 2434 bp in length
23322
23421: gap of 100 bp
23422
25550: contig of 2129 bp in length
25551
25650: gap of 100 bp
25651
28427: contig of 2577 bp in length
28428
28327: gap of 100 bp
28328
30559: contig of 2332 bp in length
30560
30659: gap of 100 bp
30660
33173: contig of 2514 bp in length
33174
33273: gap of 100 bp
33274
33374
36522: contig of 3249 bp in length
36523
36622: gap of 100 bp
40194: contig of 3572 bp in length
40195
40294: gap of 100 bp
40395
44592: contig of 5298 bp in length
44593
44692: gap of 100 bp
45693
45673: contig of 7581 bp in length
53374
53373: gap of 100 bp
53374
59245: contig of 5872 bp in length
59246
59345: gap of 100 bp
59346
56502: contig of 7157 bp in length
66503
66602: gap of 100 bp
66603
74946: contig of 8344 bp in length
74947
75047
82285: contig of 7239 bp in length
82286
82385: gap of 100 bp
82386
90377: contig of 7992 bp in length
90378
90477: gap of 100 bp
90478
99467: contig of 8990 bp in length
99468
99567: gap of 100 bp
99568
110814: contig of 11247 bp in length
110815
110914: gap of 100 bp
110915
123516: contig of 12602 bp in length
123517
123516: gap of 100 bp
123517
136720: contig of 13104 bp in length
136721
136820: gap of 100 bp
136821
142126: contig of 5306 bp in length.

FEATURES
source
1.142126
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6"
/clone="RP11-278J20"
/clone_11b="RP11 Human Male BAC"
1.16449
/note="assembly_fragment
vector_side:left"
clone_end:SP6
16550..17837
/note="assembly_fragment"
17938..19143
/note="assembly_fragment"
19244..20787
/note="assembly_fragment"
20888..23321
/note="assembly_fragment"
23422..25550
/note="assembly_fragment"
25651..28227
/note="assembly_fragment"
28328..30559

```

```
misc_feature      1..16449
                  /note="assembly_fragment
                  clone_end:SP6
                  vector_side:left"
misc_feature      16550..17837
                  /note="assembly_fragment"
misc_feature      17938..19143
                  /note="assembly_fragment"
misc_feature      19244..20787
                  /note="assembly_fragment"
misc_feature      20888..23321
                  /note="assembly_fragment"
misc_feature      23422..25550
                  /note="assembly_fragment"
misc_feature      25651..28227
                  /note="assembly_fragment"
misc_feature      28328..30559
                  /note="assembly_fragment"
```

```

misc_feature /note="assembly_fragment"
30660..33173
misc_feature /note="assembly_fragment"
33274..36522
misc_feature /note="assembly_fragment"
36623..40194
misc_feature /note="assembly_fragment"
40295..45592
misc_feature /note="assembly_fragment"
45693..53273
misc_feature /note="assembly_fragment"
53374..59245
misc_feature /note="assembly_fragment"
59346..66502
misc_feature /note="assembly_fragment"
66603..74946
misc_feature /note="assembly_fragment"
75047..82285
misc_feature /note="assembly_fragment"
82386..90377
misc_feature /note="assembly_fragment"
90478..99467
misc_feature /note="assembly_fragment"
99568..110814
misc_feature /note="assembly_fragment"
110915..112516
misc_feature /note="assembly_fragment"
123617..136720
misc_feature /note="assembly_fragment"
136821..142126
misc_feature /note="assembly_fragment"
clone_end:17
vector_side:right"

```

ORIGIN

```

Query Match      24.1%  Score 21;  DB 2;  Length 142126;
Best Local Similarity 100.0%;  Pred. No. 0.18;
Matches 21;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

QY      51 GGAGGCAAGAGACACTCTGG 71
Db      135317 GGAGGCAAGAGACACTCTGG 135337

```

```

RESULT 12
AC025018/c      159942 bp      DNA      linear      PRI 01-MAY-2002
LOCUS      AC025018      Homo sapiens chromosome 18, clone RP11-320C19, complete sequence.
DEFINITION      AC025018
AC025018      AC025018.6  GI:20377056
VERSION      HTG.
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 159942)
AUTHORS      Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL      Unpublished
TITLE      2 (bases 1 to 159942)
REFERENCE
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collimore,A., Cooke,P., Deatellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howard,J.C., Iliev,I., Johnson,R., Jones,C., Kam,L., Karakas,A.,
Klein,J., Liu,C., Liu,G., Locke,K., Landers,T., Lehoczy,J.,
Levine,R., Liew,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
Melgrim,J., Menus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,

```

```

TITLE      Direct Submission
JOURNAL    Submitted (03-MAR-2000) Whitehead Institute/MIT Center for Genome
REFERENCE  Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS    3 (bases 1 to 159942)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collimore,A.,
Cooke,A., Cooke,P., Deatellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gind,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karakas,A., Kells,C., Labocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
McCarthy,M., McEwan,P., McKernan,K., Melgrim,J., Menus,L.,
McGurk,A., MacDonald,P., Major,J., Marquis,N., Matthews,C.,
Mihova,T., Mlenga,V., Morrow,J., Naylor,T., Nguyen,C., Nicol,R.,
Nobu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,D., Peterson,C.H., Phunkhang,P., Piere,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Riese,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thoman,N., Stojanovic,N.,
Strauss,N., Sudramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trifilio,D., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

```

```

TITLE      Direct Submission
JOURNAL    Submitted (28-APR-2002) Whitehead Institute/MIT Center for Genome
REFERENCE  Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS    4 (bases 1 to 159942)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collimore,A.,
Cooke,A., Cooke,P., Deatellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gind,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karakas,A., Kells,C., Labocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
McCarthy,M., McEwan,P., McKernan,K., Melgrim,J., Menus,L.,
McGurk,A., MacDonald,P., Major,J., Marquis,N., Matthews,C.,
Mihova,T., Mlenga,V., Morrow,J., Naylor,T., Nguyen,C., Nicol,R.,
Nobu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,D., Peterson,C.H., Phunkhang,P., Piere,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Riese,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thoman,N., Stojanovic,N.,
Strauss,N., Sudramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trifilio,D., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

```

```

TITLE      Direct Submission
JOURNAL    Submitted (01-MAY-2002) Whitehead Institute/MIT Center for Genome
REFERENCE  Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS    On May 1, 2002 this sequence version replaced gi:20336169.
COMMENT    All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/XM/RepeatMasker.html
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information

```

Center project name: L7797
Center clone name: 320_C_19

| FEATURES | Location/Qualifiers |
|---------------|------------------------------------|
| source | 1..159942 |
| | /organism="Homo sapiens" |
| | /mol_type="genomic DNA" |
| | /db_xref="taxon:9606" |
| | /chromosome="18" |
| | /map="18" |
| | /clone="RP11-320C19" |
| | /clone_id="RPCT-11 Human Male BAC" |
| repeat_region | 63..203 |
| | /rpt_family="(GGA)n" |
| repeat_region | complement(733..898) |
| | /rpt_family="Char11e8" |
| repeat_region | 2441..3030 |
| | /rpt_family="LIM3c" |
| repeat_region | 3037..3241 |
| | /rpt_family="AluJo" |
| repeat_region | 3243..3813 |
| | /rpt_family="L1" |
| repeat_region | 3814..4126 |
| | /rpt_family="AluSx" |
| repeat_region | 4127..4342 |
| | /rpt_family="L1" |
| repeat_region | 4343..4545 |
| | /rpt_family="MER1B" |
| repeat_region | complement(4546..4852) |
| | /rpt_family="AluSg" |
| repeat_region | 4853..5013 |
| | /rpt_family="MER1B" |
| repeat_region | 5014..5173 |
| | /rpt_family="L1" |
| repeat_region | 5186..8358 |
| | /rpt_family="L1P3" |
| repeat_region | complement(8357..8409) |
| | /rpt_family="L1P8" |
| repeat_region | complement(8410..8556) |
| | /rpt_family="L1P" |
| repeat_region | complement(8535..8801) |
| | /rpt_family="L1P" |
| repeat_region | 8803..9894 |
| | /rpt_family="L1P8" |
| repeat_region | 9896..10358 |
| | /rpt_family="L1P8" |
| repeat_region | 10364..12455 |
| | /rpt_family="L1" |
| repeat_region | 12456..12747 |
| | /rpt_family="AluJo" |
| repeat_region | 12762..13630 |
| | /rpt_family="(TR)n" |
| unsure | complement(13146..13422) |
| | /note="single clone coverage" |
| repeat_region | 13615..14094 |
| | /rpt_family="(TATA)n" |
| repeat_region | 14053..14219 |
| | /rpt_family="(TR)n" |
| repeat_region | 14241..14392 |
| | /rpt_family="L1MA1" |
| repeat_region | 14483..14990 |
| | /rpt_family="L1MA9" |
| repeat_region | complement(15265..15543) |
| | /rpt_family="AluY" |
| repeat_region | 16050..16079 |
| | /rpt_family="(GA)n" |
| repeat_region | complement(16838..17311) |
| | /rpt_family="AluSc" |
| repeat_region | 18002..18044 |
| | /rpt_family="AT_rich" |
| repeat_region | 18259..18476 |
| | /rpt_family="MIR" |
| repeat_region | 18512..20673 |

| | |
|---------------|--------------------------|
| repeat_region | /rpt_family="L1MA1" |
| | complement(20738..20955) |
| | /rpt_family="MER56A" |
| repeat_region | 21395..21417 |
| | /rpt_family="AT_rich" |
| repeat_region | 21958..22317 |
| | /rpt_family="THE1B" |
| repeat_region | 22315..23806 |
| | /rpt_family="THE1B-int" |
| repeat_region | 23807..24173 |
| | /rpt_family="THE1B" |
| repeat_region | 24189..24239 |
| | /rpt_family="(TCTG)n" |
| repeat_region | 24244..24281 |
| | /rpt_family="(CA)n" |
| repeat_region | 24437..24468 |
| | /rpt_family="AT_rich" |
| repeat_region | 25078..25307 |
| | /rpt_family="MIR" |
| repeat_region | 25413..25569 |
| | /rpt_family="FLAM_A" |
| repeat_region | 26749..26790 |
| | /rpt_family="AT_rich" |
| repeat_region | complement(26905..27421) |
| | /rpt_family="L1MC1" |
| repeat_region | 27427..27938 |
| | /rpt_family="MER6B" |

Query Match 24.1% Score 21 DB 9 Length 159942;
Best Local Similarity 100.0% Pred. No. 0.18;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GGAGGCAAGAGCACTCTGG 71
DQ 91013 GGAGGCAAGAGCACTCTGG 90993

| RESULT 13 | AL512427 | LOCUS | DEFINITION | ACCESSION | VERSION | KEYWORDS | SOURCE | ORGANISM | REFERENCE | AUTHORS | TITLE | JOURNAL | COMMENT | | | | |
|-----------|-----------|-------|--|-----------|-------------|-------------|--------|----------------------|-------------------------|--|-----------------------------|--|--|-------------------------|--|-----------------------------|--|
| AL512427 | 201657 bp | DNA | Human DNA sequence from clone RP11-325M4 on chromosome 6, complete sequence. | AL512427 | AL512427.10 | GI:14270153 | HTG. | Homo sapiens (human) | Submitted (25-MAY-2001) | Sanger Centre, Hinxton, Cambridgeshire, UK | Request: clones@anger.ac.uk | On May 31, 2001 this sequence version replaced GI:13235016. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. | CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk | Submitted (25-MAY-2001) | Sanger Centre, Hinxton, Cambridgeshire, UK | Request: clones@anger.ac.uk | On May 31, 2001 this sequence version replaced GI:13235016. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. |

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMBPP; Information on the WORMBPP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormp This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6>
RP11-325M4 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBACE3.6

This sequence is the entire insert of clone RP11-325M4.

FEATURES

```
source
location/Qualifiers
1..201657
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="6"
  /clone_lib="RP11-325M4"
  /clone_1ib="RP11-11.2"
  7..696
    /note="match: GSS: Em:AQ536961"
  14..725
    /note="match: GSS: Em:AQ395017"
  19..429
    /note="match: GSS: Em:AQ390879"
  157..2551
    /note="TRIGER1 repeat: matches 11..2418 of consensus"
  2582..2665
    /note="42 copies 2 mer ta 63% conserved"
  3974..4332
    /note="LTR1A1 repeat: matches 1..363 of consensus"
  4333..4378
    /note="23 copies 2 mer gc 82% conserved"
  /note="match: GSS: Em:AQ455938"
  /complement(4827..5207)
  /note="match: GSS: Em:AQ375389"
  /complement(4960..5519)
  /note="match: GSS: Em:AQ377395"
  /complement(4963..5512)
  /note="match: GSS: Em:AQ537274"
  /complement(5221..5487)
  /note="match: GSS: Em:AQ525186"
  5519..6021
    /note="match: GSS: Em:AQ525186"
  6849..8904
    /note="MIR repeat: matches 196..252 of consensus"
  8926..9202
    /note="LIPAS repeat: matches 5866..6142 of consensus"
  9203..9537
    /note="LIPAS repeat: matches 5536..5873 of consensus"
  9553..9709
    /note="MIR repeat: matches 2..155 of consensus"
  10469..10786
    /note="AluJb repeat: matches 1..306 of consensus"
  12502..12885
    /note="LTR1B repeat: matches 1..390 of consensus"
  13350..13494
    /note="MIR repeat: matches 6..154 of consensus"
  /complement(13695..14217)
  /note="match: GSS: Em:AQ732279"
  /complement(13702..14228)
  /note="match: GSS: Em:AQ892616"
  /complement(13736..14178)
  /note="match: GSS: Em:AQ571323"
  /complement(13743..14205)
  /note="match: GSS: Em:AQ808008"
  14570..14850
    /note="match: GSS: Em:AQ192455"
  15424..15596
    /note="LIM1 repeat: matches -1388..-1207 of consensus"
  15639..15710
    /note="AluJd/FRAM repeat: matches 223..294 of consensus"
  15713..16295
    /note="LIR9 repeat: matches 14..625 of consensus"
  16328..16468
    /note="AluJb repeat: matches 1..142 of consensus"
  16469..19085
    /note="LIM1 repeat: matches -1208..-1472 of consensus"
  19085..19268
    /note="LIM3 repeat: matches 2410..2592 of consensus"
  19269..20303
    /note="MER11C repeat: matches 1..1071 of consensus"
  20304..20857
    /note="LIM3 repeat: matches 2592..3144 of consensus"
  20858..21128
    /note="AluJd repeat: matches 21..283 of consensus"
  21129..23232
    /note="LIM3 repeat: matches 3144..5258 of consensus"
  23233..23399
    /note="LIM3 repeat: matches 5413..5567 of consensus"
  23400..23651
    /note="AluJc repeat: matches 56..307 of consensus"
  23652..24315
    /note="LIM3 repeat: matches 5567..6304 of consensus"
  24316..24336
    /note="LIM3 repeat: matches 5258..5277 of consensus"
  24522..24587
    /note="33 copies 2 mer aa 69% conserved"
  25491..25845
    /note="match: GSS: Em:AQ59990"
  26640..26689
    /note="25 copies 2 mer ga 84% conserved"
  27322..29166
    /note="LIP14 repeat: matches 4003..5855 of consensus"
  29165..29223
    /note="LIP repeat: matches 3992..4045 of consensus"
  29227..29515
    /note="LIPAS repeat: matches 5859..6163 of consensus"
  31308..31392
    /note="LTR1H repeat: matches 72..156 of consensus"
  31456..31724
    /note="LTR1H repeat: matches 247..532 of consensus"
  31857..32276
    /note="MSRB repeat: matches 1..426 of consensus"
  32629..33073
    /note="match: GSS: Em:AQ695871"
  33032..331923
    /note="LIMB3 repeat: matches 5290..6182 of consensus"
  34754..34858
    /note="MIR repeat: matches 65..179 of consensus"
  35312..35730
    /note="MST1 repeat: matches 3..426 of consensus"
  /complement(36413..36996)
  /note="match: GSS: Em:AQ399585"
  /complement(36763..37094)
  /note="match: GSS: Em:B39390"
  36997..37476
    /note="match: GSS: Em:AQ719307"
  37484..37907
    /note="MSTRB repeat: matches 1..426 of consensus"
  39019..39197
    /note="MSTR20 repeat: matches 4..177 of consensus"
  39315..39409
    /note="L2 repeat: matches 2629..2748 of consensus"
  39598..41702
    /note="LIPAS repeat: matches 4040..6151 of consensus"
  41930..41994
    /note="L2 repeat: matches 2624..2650 of consensus"
  41938..42385
    /note="match: GSS: Em:AQ240510"
  41938..42298
    /note="match: GSS: Em:B31098"
  41956..42296
    /note="match: GSS: Em:B31099"
  /complement(42138..42424)
  /note="match: GSS: Em:AQ527813"
  42335..42614
    /note="AluJd repeat: matches 1..282 of consensus"
```

```

repeat_region      42751..42826
                    /note="MER5B repeat: matches 2..78 of consensus"
repeat_region      42835..42896
                    /note="MER92C repeat: matches 494..546 of consensus"
repeat_region      42880..43251
                    /note="MER67D repeat: matches 71..436 of consensus"
repeat_region      43252..44618
                    /note="LIP2A8 repeat: matches 2430..3809 of consensus"
repeat_region      44610..46944
                    /note="LIP2A8 repeat: matches 3824..6163 of consensus"
repeat_region      46945..46980
                    /note="18 copies 2 mar aa 80% conserved"
repeat_region      46983..47063
                    /note="MER67D repeat: matches 5..88 of consensus"
repeat_region      47347..47404
                    /note="MER53 repeat: matches 1..58 of consensus"
repeat_region      49755..50074
                    /note="Alu10 repeat: matches 1..312 of consensus"
repeat_region      51520..51763
                    /note="LIR29 repeat: matches 423..619 of consensus"

```

Query Match 24.1% Score 21; DB 9; Length 201657;

Best Local Similarity 100.0%; Pred. No. 0.18; Mismatches 0; Indels 0; Gaps 0;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 GGAGGCAAGACACTCTGG 71

Db 96897 GGAGGCAAGACACTCTGG 96917

```

RESULT 14
LOCUS      AF245511/c      2576 bp      mRNA      linear      VRT 08-AUG-2001
DEFINITION Gallus gallus ubiquitin protein ligase mRNA, partial cds.
ACCESSION  AF245511 AF251045
VERSION     AF245511.2 GI:15088510
KEYWORDS
SOURCE      Gallus gallus (chicken)
ORGANISM    Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 2576)
AUTHORS    Lomax,M.I., Huang,L., Cho,Y., Gong,T.L. and Altschuler,R.A.
TITLE      Differential display and gene arrays to examine auditory plasticity
JOURNAL    Hear. Res. 147 (1-2), 293-302 (2000)
MEDLINE    20420030
PUBMED     10962193
REFERENCE   2 (bases 1 to 2576)
AUTHORS    Lomax,M.I., Huang,L., Cho,Y., Gong,T.L. and Altschuler,R.A.
TITLE      Direct Submission
JOURNAL    Submitted (15-MAR-2000) Otolaryngology, University of Michigan,
            1150 W. Med. Ctr. Dr., Ann Arbor, MI 48109, USA
            3 (bases 1 to 2576)
AUTHORS    Lomax,M.I., Huang,L., Cho,Y., Gong,T.L. and Altschuler,R.A.
TITLE      Direct Submission
JOURNAL    Submitted (06-AUG-2001) Otolaryngology, University of Michigan,
            1150 W. Med. Ctr. Dr., Ann Arbor, MI 48109, USA
REMARK     Sequence update by submitter
COMMENT    On or before Aug 8, 2001 this sequence version replaced
            GI:1507056, GI:12382234.
FEATURES   Location/Qualifiers
            source
            1..2576
                /organism="Gallus gallus"
                /mol_type="mRNA"
                /db_xref="taxon:9031"
                /tissue_type="inner ear"
                <1..1201
                /note="UBB3B"
                /codon_start=2
                /product="ubiquitin protein ligase"
                /protein_id="AA653076.2"
                /db_xref="GI:15088511"

```

```

/translation="LVETSSASPHYHTITIRSRMLEDGYEQRLQSGNANKGYLRVK
FYNDLGVBDAGIDODGVFKPELEELIKVFPDNLVFTSGDERLYSPSYHENV
LQLEFVKGMLGKAYREGIVDVPASPELSQLGHHHSVFPYSVDLPSLDSEYKN
LTSIKRYDSDIDDLGLTISYEDVNGQLVCHETLVGGCTTTPYENKTSYHIANHFR
MHTQIKSOTPAULISGRSITIKPEWTRMFSAPLQRLISGDNALDELQKHTYYGG
FGSHRVITLMDIILANDFSPEERAMFQVYCSRPPLGAYIKPFPFSJRCVHSD
DDDTGDTLGSVLRGFTTRKKEPGGRLPTSSCFMLKLPYNSKXSILREKLYAISM
NTGFELS"
1202..2547

```

ORIGIN 3'UTR

Query Match 23.0% Score 20; DB 5; Length 2576;

Best Local Similarity 100.0%; Pred. No. 1.2; Mismatches 0; Indels 0; Gaps 0;

Qy 13 TTATTGAAGTGAGATTTT 32

Db 495 TTATTGAAGTGAGATTTT 476

```

RESULT 15
LOCUS      AC008708/c      149378 bp      DNA      linear      PRI 25-JUN-2002
DEFINITION Homo sapiens chromosome 5 clone CTB-78F1, complete sequence.
ACCESSION  AC008708
VERSION     AC008708.6 GI:21553203
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 149378)
AUTHORS    Doe Joint Genome Institute and Stanford Human Genome Center.
TITLE      Direct Submission
JOURNAL    Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            2 (bases 1 to 149378)
AUTHORS    Doe Joint Genome Institute.
TITLE      Direct Submission
JOURNAL    Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            3 (bases 1 to 149378)
AUTHORS    Doe Joint Genome Institute and Stanford Human Genome Center.
TITLE      Direct Submission
JOURNAL    Submitted (25-JUN-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
            On Jun 25, 2002 this sequence version replaced GI:19224700.
            www.jgi.doe.gov
            www.sngc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
NOTE: BACTERIAL TRANSPOSON excised at 28650.
FEATURES   Location/Qualifiers
            source
            1..149378
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="5"
                /clone="CTB-78F1"

```

ORIGIN

Query Match 23.0% Score 20; DB 9; Length 149378;

Best Local Similarity 100.0%; Pred. No. 0.72; Mismatches 0; Indels 0; Gaps 0;

Qy 21 AGTAGATTTCCTGGCCTG 40

|||||

Db 15534 AGTGAGATTCTGTGACCTG 15515

RESULT 16
AC103355/c
LOCUS AC103355
DEFINITION Mus musculus chromosome 14 clone RP24-482C10 map 14, *** SEQUENCING
IN PROGRESS ***, 3 unordered pieces.
ACCESSION AC103355
VERSION AC103355.8 GI:39725783
KEYWORDS HTG: HTGS PHASE1; HTGS_FULFILL; HTGS_ACTIVERIN.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 195708)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus chromosome 14, clone RP24-482C10
2 (bases 1 to 195708)
Unpublished
Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckgalter, B.,
Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, N., Collins, S., Collymore, A., Cook, A.,
Cooke, P., Dearellano, K., Dewar, K., Diaz, J., Dodge, S., Fero, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K.,
Lamazaras, R., Lander, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McGowan, P., McKernan, K., McPeckers, R., Meldrum, J.,
Meneas, L., Minova, T., Menga, V., Murphy, T., Naylor, J., Nguyen, C.,
Nord, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travers, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (24-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 195708)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Bouckgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J., S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hachopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lander, E., Levine, R.,
Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Meneas, L., Minova, T., Menga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Nord, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (11-DEC-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 11, 2003 this sequence version replaced gi:38454415.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center Project name: L16635
Center Clone name: 482_C_10

* NOTES: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 146464: contig of 146464 bp in length
* 146465 146564: gap of 100 bp
* 146565 192519: contig of 45955 bp in length
* 192520 192619: gap of 100 bp
* 192620 195708: contig of 3089 bp in length.
Location/Qualifiers
1. 195708
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="14"
/map="14"
/clone="RP24-482C10"
/clone_1b="PFC1-24 Male Mouse BAC"
ORIGIN
Query Match 23.0%; Score 20; DB 2; Length 195708;
Best Local Similarity 100.0%; Fred. No. 0.7;
Matches 20; Conservative 0; Mismatches 0; Gaps 0;
Gy 32 TCTGGCCTGAGACTTGACG 51
Db 134902 TCTGGCCTGAGACTTGACG 134883
RESULT 17
AC073278/c
LOCUS AC073278 200807 bp DNA linear PRI 08-OCT-2003
DEFINITION Homo sapiens BAC clone RP11-66118 from 7, complete sequence.
ACCESSION AC073278
VERSION AC073278.9 GI:15145619
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 200807)
Toward a complete human genome sequence
Sulston, J.E. and Wilson, R.
2 (bases 1 to 200807)
9847074
PUBMED 99063792
MEDLINE 99063792
Genome Res. 8 (11), 1097-1108 (1998)
REFERENCE
1. (bases 1 to 200807)
Toward a complete human genome sequence
Sulston, J.E. and Wilson, R.
2 (bases 1 to 200807)
9847074
PUBMED 99063792
MEDLINE 99063792
Genome Res. 8 (11), 1097-1108 (1998)
REFERENCE
Abbott, A., Boyer, E., Elliott, G. and Harris, A.
The sequence of Homo sapiens BAC clone RP11-66118
Unpublished (2001)
3 (bases 1 to 200807)
REFERENCE
Waterson, R.H.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 200807)
REFERENCE
Waterson, R.H.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-2001) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
5 (bases 1 to 200807)
AUTHORS
Waterson, R.
JOURNAL
Submitted (09-JUN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 200807)
REFERENCE
AUTHORS
Wilson, R.
TITLE
Direct Submission
JOURNAL
Submitted (08-OCT-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Aug 9, 2001 this sequence version replaced gi:14476026.
COMMENT

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: sapienewatson.wustl.edu

Summary Statistics
Center project name: H_NH0661108

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
Project (Eric D. Green, Director), John D. McPherson in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
<http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send
<mailto:regreen@nhgri.nih.gov>, or see <http://genome.wustl.edu>

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa, K., Moon, P. Y., Zhao, B., Frenken, E.,
Tateno, M., Catanese, J. J., and de Jong, P. J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
and coworkers at <http://www.chori.org>
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-627B12 the clone sequenced
to the right is RP5-1102A12, 2000 bp overlap. Actual start of this
clone is at base position 1 of RP11-66118.

FEATURES

source
1. 200807
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/map="7"
/clone="RP11-66118"
/clone_lib="RPCI-11"
38..98
repeat_region
/rpt_family="CRI"
645..935
repeat_region
/rpt_family="Alu"
1113..1410
repeat_region

repeat_region /rpt_family="Alu"
1858..2040
/rpt_family="Alu"
repeat_region /rpt_family="Alu"
2202..2308
/rpt_family="Alu"
repeat_region /rpt_family="Alu"
2392..2428
/rpt_family="L2"
2615..3035
/rpt_family="Alu"
3036..3197
/rpt_family="MIR"
3357..3458
/rpt_family="C-rich"
3456..3785
/rpt_family="Alu"
3982..4295
/rpt_family="Alu"
4302..4377
/rpt_family="L1"
4518..4790
/rpt_family="Alu"
4791..4812
/rpt_family="CA)n"
4852..4934
/rpt_family="L1"
4927..5144
/rpt_family="L1"
5145..5427
/rpt_family="Alu"
5428..5796
/rpt_family="L1"
5882..5915
/rpt_family="Alu"
6061..6383
/rpt_family="L1"
6384..6430
/rpt_family="TG)n"
6431..6684
/rpt_family="Alu"
6685..6723
/rpt_family="L1"
6795..6825
/rpt_family="CA)n"
6826..6920
/rpt_family="TA)n"
6932..7226
/rpt_family="Alu"
7245..7370
/rpt_family="Alu"
7574..7599
/rpt_family="AT-rich"
7868..8008
/rpt_family="Alu"
8031..8053
/rpt_family="BC200"
8054..8418
/rpt_family="MIR"
8419..8543
/rpt_family="BC200"
8671..8796
/rpt_family="MIR"
9175..9453
/rpt_family="Alu"
9481..9767
/rpt_family="Alu"
9772..9899
/rpt_family="Alu"
9900..9936
/rpt_family="CA)n"
9937..10100
/rpt_family="Alu"
10413..10450
/rpt_family="TTCTC)n"


```

repeat_region      10451..10753
                    /rpt_family="Alu"
repeat_region      11141..11472
                    /rpt_family="L2"
repeat_region      11473..11745
                    /rpt_family="Alu"
repeat_region      11746..11965
                    /rpt_family="L2"
repeat_region      12550..12690
                    /rpt_family="Alu"
repeat_region      13003..13140
                    /rpt_family="L1"
repeat_region      13141..13441
                    /rpt_family="Alu"
repeat_region      13442..13692
                    /rpt_family="L1"
repeat_region      13693..13979
                    /rpt_family="Alu"
repeat_region      13980..14002
                    /rpt_family="L1"
repeat_region      14003..14312
                    /rpt_family="Alu"
repeat_region      14313..14416
                    /rpt_family="L1"
repeat_region      14559..14921

```

```

Query Match      23.0% Score 20; DB 9; Length 200807;
Best Local Similarity 100.0%; Pred. NO. 0.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      23 TGAGATTTTTCGGCCTGAG 42
Db      59425 TGAGATTTTTCGGCCTGAG 59406

```

```

RESULT 18
AC141733
LOCUS      AC141733
DEFINITION AC141733 clone CH224-58F20, WORKING DRAFT SEQUENCE, 71
unordered pieces.
ACCESSION AC141733
VERSION AC141733.1 GI:29123917
KEYWORDS HTGS PHASE1; HTGS DRAFT.
SOURCE HTGS mellifera (honeybee)
ORGANISM Apis mellifera
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
          Apidae; Apis.
          1 (bases 1 to 223920)
Muzny,D.M., Adams,C., Adio-Oduola,B., All-oman,F.R., Allen,C.,
Albrooks,S.L., Amaral-Tunget,H.C., Are,J.R., Ayala,M., Banks,T.,
Barbieri,J., Benton,U., Bimaga,K., Blankenburg,K., Bonin,D.,
Bouck,J., Bowls,S., Brileva,M., Brown,M., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Doutwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Fall,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Fratcz,P.,
Gabriel,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.B.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S.,
Karlsoson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louieged,H.,
Lozago,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

Maesey,E., Mawney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokwenkwo,S., Ogih,M., Okumura,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojudoan,I., Rolle,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoshitari,N., Slason,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Swatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,R., Thomas,S.,
Uman,K., Vaequez,L., Vera,V., Villanov,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleczky,R., Woodson,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 223920)
Worley,K.C.
Submitted (19-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
-----
Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: AMER
Center clone name: CH224-58F20
-----
Summary Statistics
Sequencing vector: Plasmid,
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 208891 bases at least Q40
Consensus quality: 217674 bases at least Q30
Consensus quality: 223660 bases at least Q20
Estimated insert size: 203635; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 71 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1287 1286: contig of 1286 bp in length
1387 1386: gap of unknown length
2587 2587: contig of 1201 bp in length
2588 2587: gap of unknown length
2687 2687: gap of unknown length
3873 3873: contig of 1186 bp in length
3874 3873: gap of unknown length
3973 3973: gap of unknown length
5300 5300: contig of 1327 bp in length
5301 5300: gap of unknown length
5400 5400: gap of unknown length
6899 6899: contig of 1499 bp in length
6900 6899: gap of unknown length
7000 6998: contig of 1499 bp in length
8498 8498: gap of unknown length
8499 8498: gap of unknown length
8599 8598: contig of 1235 bp in length
9834 9833: gap of unknown length
9834 9833: gap of unknown length
9934 9933: contig of 1152 bp in length
11086 11085: contig of 1457 bp in length
11186 11185: gap of unknown length
12643 12642: gap of unknown length
12743 12742: contig of 1197 bp in length
13939 13938: contig of 1197 bp in length
13940 13939: gap of unknown length
14039 14038: gap of unknown length
15762 15761: contig of 1723 bp in length

```

15763 15862: gap of unknown length
 * 15863 17042: contig of 1180 bp in length
 * 17043 17142: gap of unknown length
 * 17143 18188: contig of 1046 bp in length
 * 18189 18288: gap of unknown length
 * 18289 19305: contig of 1017 bp in length
 * 19306 19405: gap of unknown length
 * 19406 20891: contig of 1486 bp in length
 * 20892 20992: gap of unknown length
 * 20993 22172: contig of 1181 bp in length
 * 22173 22272: gap of unknown length
 * 22273 23607: contig of 1335 bp in length
 * 23608 23707: gap of unknown length
 * 23708 25369: contig of 1662 bp in length
 * 25370 25469: gap of unknown length
 * 25470 26735: contig of 1266 bp in length
 * 26736 26835: gap of unknown length
 * 26836 28099: contig of 1264 bp in length
 * 28100 28200: gap of unknown length
 * 28201 29320: contig of 1021 bp in length
 * 29321 31314: gap of unknown length
 * 31315 31414: gap of unknown length
 * 31415 32562: contig of 1148 bp in length
 * 32563 32662: gap of unknown length
 * 32663 34125: contig of 1463 bp in length
 * 34126 34225: gap of unknown length
 * 34226 35896: contig of 1671 bp in length
 * 35897 35966: gap of unknown length
 * 35967 37348: contig of 1352 bp in length
 * 37349 37448: gap of unknown length
 * 37450 39365: contig of 1917 bp in length
 * 39366 39465: gap of unknown length
 * 39466 40958: contig of 1493 bp in length
 * 40959 41058: gap of unknown length
 * 41059 42440: contig of 1382 bp in length
 * 42441 42540: gap of unknown length
 * 42541 43967: contig of 1427 bp in length
 * 43968 44067: gap of unknown length
 * 44068 45739: contig of 1672 bp in length
 * 45740 45839: gap of unknown length
 * 45840 47140: contig of 1301 bp in length
 * 47141 47240: gap of unknown length
 * 47241 48958: contig of 1718 bp in length
 * 48959 49058: gap of unknown length
 * 49059 51365: contig of 2307 bp in length
 * 51366 51465: gap of unknown length
 * 51466 53056: contig of 1591 bp in length
 * 53057 53156: gap of unknown length
 * 53157 54710: contig of 1554 bp in length
 * 54711 54810: gap of unknown length
 * 54811 56040: contig of 1230 bp in length
 * 56041 56140: gap of unknown length
 * 56141 58307: contig of 2167 bp in length
 * 58308 58407: gap of unknown length
 * 58408 60071: contig of 1664 bp in length
 * 60072 60171: gap of unknown length
 * 60172 61494: contig of 1323 bp in length
 * 61495 61594: gap of unknown length
 * 61595 63574: contig of 1980 bp in length
 * 63575 63674: gap of unknown length
 * 63675 64825: contig of 1151 bp in length
 * 64826 64925: gap of unknown length
 * 64927 66277: contig of 1352 bp in length
 * 66278 66377: gap of unknown length
 * 66378 68973: gap of 2596 bp in length
 * 68974 69073: gap of unknown length
 * 69074 71541: contig of 2466 bp in length
 * 71542 71641: gap of unknown length
 * 71642 73546: contig of 1905 bp in length
 * 73547 73646: gap of unknown length
 * 73647 75530: contig of 1884 bp in length
 * 75531 75630: gap of unknown length

75631 77877: contig of 2247 bp in length
 * 77878 77977: gap of unknown length
 * 77978 79836: contig of 1859 bp in length
 * 79837 79936: gap of unknown length
 * 79937 81907: contig of 1971 bp in length
 * 81908 82007: gap of unknown length
 * 82008 83869: contig of 1862 bp in length
 * 83870 83969: gap of unknown length
 * 83970 86241: contig of 2272 bp in length
 * 86242 86341: gap of unknown length
 * 86342 89094: contig of 2753 bp in length
 * 89095 89194: gap of unknown length
 * 89195 92079: contig of 2885 bp in length
 * 92080 92179: gap of unknown length
 * 92180 96535: contig of 4356 bp in length
 * 96536 96635: gap of unknown length
 * 96636 100264: contig of 3629 bp in length
 * 100265 100364: gap of unknown length

Query Match 23.0%; Score 20; DB 2; Length 223920;
 Best Local Similarity 100.0%; Pred. No. 0.69;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AATTGAAGTCAGATTCTCT 34
 Db 106395 AATTGAAGTCAGATTCTCT 106414

RESULT 19
 AC141730/c
 LOCUS
 DEFINITION
 AC141730 274341 bp DNA linear HTG 19-MAR-2003
 Apis mellifera clone CH224-58F21, WORKING DRAFT SEQUENCE, 90
 unordered pieces.
 AC141730
 VERSION
 AC141730.1 GI:29123914
 KEYWORDS
 HTG; HTGS PHASE1; HTGS DRAFT.
 SOURCE
 Apis mellifera (honeybee)
 ORGANISM
 Apis mellifera
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Hymenoptera; Apoecrita; Aculeata; Apoidea;
 Apidae; Apis.
 1 (bases 1 to 274341)

REFERENCE

AUTHORS

Mazny,D.M., Adams,C., Adio-Oduola,B., Ali-Isman,F.R., Allen,C.,
 Alshrocks,S.L., Anaratunge,H.C., Are,J.R., Ayele,M., Banke,T.,
 Barbara,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D.,
 Bouay,C., Bowls,S., Briteva,M., Brown,B., Brown,M., Bryant,N.P.,
 Buhray,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dim,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.U.,
 Barnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguno,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gottell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hayes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
 Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratochvic,J., Kiremitli,A., Landry,N., Leal,B., Levele,L.C., Lewis,L.,
 Li,J., Li,Z., Licharge,O., Lien,C., Liu,J., Liu,W., Louieged,H.,
 Lozada,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Mabeahwar,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,N., Newton,U., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogulu,M., Okwuonu,G.,
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Plims,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rives,M., Rojas,A., Rojudoan,I., Rolle,M., Ruiz,S., Saverly,C.,

Scherer, S., Scott, G., Shen, H., Shooshari, N., Sisson, I.,
Sodergren, E., Sotak, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Tumani, K., Vasquez, L., Vera, V., Villalon, D., Vinton, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.

Unpublished
Direct Submission
2 (bases 1 to 274341)
Mortley, K.C.

Submitted (19-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: AMED
Center clone name: CH224-58P21
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 265284 bases at least Q40
Consensus quality: 272666 bases at least Q30
Consensus quality: 278765 bases at least Q20
Estimated insert size: 242717; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 90 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1269: contig of 1269 bp in length
1 1270 1369: gap of unknown length
1 1370 2530: contig of 1161 bp in length
2531 2630: gap of unknown length
2631 3633: contig of 1003 bp in length
3634 3733: gap of unknown length
3734 4734: contig of 1001 bp in length
4735 4834: gap of unknown length
4835 6074: contig of 1240 bp in length
6075 6174: gap of unknown length
6175 7362: contig of 1188 bp in length
7363 7462: gap of unknown length
7463 8561: contig of 1099 bp in length
8562 8661: gap of unknown length
8662 10216: contig of 1555 bp in length
10217 10316: gap of unknown length
10317 11411: contig of 1095 bp in length
11412 11511: gap of unknown length
11512 13018: contig of 1507 bp in length
13019 13118: gap of unknown length
13119 14393: contig of 1275 bp in length
14394 14493: gap of unknown length
14494 15606: contig of 1113 bp in length
15607 15706: gap of unknown length
15707 16994: contig of 1288 bp in length
16995 17094: gap of unknown length
17095 18584: contig of 1490 bp in length
18585 18684: gap of unknown length

18685 19727: contig of 1043 bp in length
19728 21324: contig of 1497 bp in length
21325 21424: gap of unknown length
21425 22440: contig of 1016 bp in length
22441 22540: gap of unknown length
22541 23764: contig of 1224 bp in length
23765 23864: gap of unknown length
23865 25004: contig of 1140 bp in length
25005 25105: gap of unknown length
25106 26501: contig of 1397 bp in length
26502 26602: gap of unknown length
26603 27734: contig of 1033 bp in length
27735 28872: gap of unknown length
28873 28972: contig of 1138 bp in length
28973 30355: gap of unknown length
30356 30455: contig of 1383 bp in length
30456 31483: gap of unknown length
31484 31583: contig of 1028 bp in length
31584 32937: gap of unknown length
32938 33037: contig of 1354 bp in length
33039 34788: gap of unknown length
34789 34888: contig of 1751 bp in length
34889 36164: gap of unknown length
36165 36264: contig of 1276 bp in length
36265 37365: gap of unknown length
37366 37465: contig of 1101 bp in length
37466 38815: gap of unknown length
38816 38916: contig of 1350 bp in length
38917 40270: gap of unknown length
40271 40370: contig of 1355 bp in length
40371 41659: gap of unknown length
41660 41759: contig of 1289 bp in length
41760 43462: gap of unknown length
43463 43562: contig of 1703 bp in length
43563 44708: gap of unknown length
44709 44808: contig of 1146 bp in length
44809 46764: gap of unknown length
46765 46864: contig of 1956 bp in length
46865 48291: gap of unknown length
48292 48391: contig of 1427 bp in length
48392 49854: gap of unknown length
49855 49954: contig of 1463 bp in length
49955 51847: gap of unknown length
51848 51946: contig of 1892 bp in length
51947 54235: gap of unknown length
54236 54335: contig of 2289 bp in length
54336 55378: gap of unknown length
55379 55478: contig of 1043 bp in length
55479 57046: gap of unknown length
57047 57146: contig of 1567 bp in length
57147 58809: gap of unknown length
58810 58909: contig of 1664 bp in length
58910 60440: gap of unknown length
60441 60540: contig of 1531 bp in length
60541 62126: gap of unknown length
62127 62226: contig of 1586 bp in length
62227 63825: gap of unknown length
63826 63925: contig of 1599 bp in length
63926 65418: gap of unknown length
65419 65518: contig of 1493 bp in length
65519 66665: gap of unknown length
66666 66765: contig of 1147 bp in length
66766 68207: gap of unknown length
68208 68307: contig of 1442 bp in length
68309 69732: gap of unknown length
69733 69832: contig of 1425 bp in length
69833 71212: gap of unknown length
71213 71312: contig of 1380 bp in length
71313 72563: gap of unknown length
72564 74224: contig of 1251 bp in length
74225 74264: gap of unknown length
74265 74266: contig of 1561 bp in length

```

*      74325      74324: gap of unknown length
*      74325      75580: contig of 1256 bp in length
*      75581      75580: gap of unknown length
*      75681      77306: contig of 1526 bp in length
*      77207      77306: gap of unknown length
*      77307      78998: contig of 1692 bp in length
*      78999      79098: gap of unknown length
*      79099      80808: contig of 1710 bp in length
*      80809      80908: gap of unknown length
*      80909      84064: contig of 3156 bp in length
*      84065      84164: gap of unknown length

```

```

Query Match      23.0%; Score 20; DB 2; Length 274341;
Best Local Similarity 100.0%; Pred. No. 0.67; 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0;

```

```

QY      15 AATTGAAGTGAGATTCTTCT 34
        |||||
Db      65369 AATTGAAGTGAGATTCTTCT 65350

```

Search completed: March 13, 2004, 08:36:36
 Job time : 2090 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2004, 00:39:18 / Search time 302 seconds
(without alignments)
1223.819 Million cell updates/sec

Title: US-09-966-880A-11

Perfect score: 87
Sequence: 1 agagaacacatcatatga.....ctgacacacatcatgacag 87

Scoring table: OLIGO_NTC
Gapop 60.0, Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size: 20

Total number of hits satisfying chosen parameters: 10

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database:

N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 87 | 100.0 | 87 | 3 AAC55315 | AAC55315 Human act |
| 2 | 87 | 100.0 | 2818 | 3 AAC55312 | AAC55312 Human act |
| 3 | 87 | 100.0 | 5514 | 3 AAC55313 | AAC55313 Human act |
| 4 | 87 | 100.0 | 11204 | 3 AAC55339 | AAC55339 Human act |
| 5 | 87 | 100.0 | 11204 | 6 ABS73286 | ABS73286 DNA encod |
| 6 | 84 | 96.6 | 1543 | 7 ABX05468 | ABX05468 Human nov |
| 7 | 84 | 96.6 | 2791 | 6 ABS73287 | ABS73287 DNA encod |
| 8 | 84 | 96.6 | 2791 | 6 ABS73288 | ABS73288 DNA encod |
| 9 | 26 | 29.9 | 26 | 3 AAC55335 | AAC55335 Human act |
| 10 | 20 | 23.0 | 20 | 8 ADB88945 | ADB88945 Activatio |

ALIGNMENTS

RESULT 1
AAC55315
XX AAC55315 standard; DNA; 87 BP.
XX
AC AAC55315;
XX
DT 05-FEB-2001 (first entry)
XX

DE Human activation-induced cytidine deaminase exon 1 SEQ ID NO:11.
KW Activation-induced cytidine deaminase; AID; cytidine deaminase;
KW immune related disease; allergy; allergic disease; anti-allergic;
KW antineoplastic; antiasthmatic; ophthalmological; anti-HIV; dermatological;
KW gene therapy; B cell associated immune system disorder; food allergy;
KW immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
KW IGA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis;
KW drug allergy; allergic rhinitis; Rosen disease; Digeorge disease; AIDS;
KW ataxia telangiectasia; common variable immunodeficiency disorder;
KW major histocompatibility class II deficiency disease;
KW auto immunodeficiency syndrome; Igg subclass selection disorder; ds.

OS Homo sapiens.

PN WO200058480-A1.

PD 05-OCT-2000.

PF 28-MAR-2000; 2000WO-JP001918.

PR 29-MAR-1999; 99JP-00087192.

PR 24-JUN-1999; 99JP-00178999.

PR 27-DEC-1999; 99JP-00371382.

PA (NTSB) JAPAN TOBACCO INC.

PI (HONJ/) HONJO T.

PI Honjo T, Muramatsu M;

DR WPI, 2000-611715/58.

PT Nucleic acid encoding activation induced cytidine deaminase, useful as a

PS target for drug development for immune-related diseases including

PS allergies.

PS Claim 18; Page 150; 174pp; Japanese.

CC The present invention describes an activation-induced cytidine deaminase
CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
CC cytidine activity similar to APOBEC-1. AID has anti-allergic, antineoplastic,
CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
CC and can be used in gene therapy. AID polynucleotides are useful in
CC methods for identifying drugs for the treatment of B cell associated
CC immune system disorders, immunodeficiency diseases and allergies, such as
CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
CC globulinaemia, atopic dermatitis, allergic colitis, asthma, food allergy,
CC drug allergy, allergic rhinitis, Rosen disease, Digeorge disease, ataxia
CC telangiectasia, common variable immunodeficiency disorder, MHC (major
CC histocompatibility class) class II deficiency disease, AIDS (auto
CC immunodeficiency syndrome), elevated Igg disorder, and Igg subclass
CC immunodeficiency disorder. The DNA sequences encoding AID may be used for gene
CC therapy and the antibodies to the AID protein may be used for diagnosis
CC and treatment of these disorders. The present sequence represents the
CC exon 1 genomic DNA sequence of human AID

CC Sequence 87 BP; 28 A; 17 C; 23 G; 19 T; 0 U; 0 Other;

CC Query Match 100.0%; Score 87; DB 3; Length 87;

CC Best local similarity 100.0%; Pred. No. 8.6e-37; Indels 0; Gaps 0;

CC Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGACCATCATTAATTAAGTGAATTTTCTGCGCTGAGACTTGACAGGAGCAAGA 60

DB 1 AGAGACCATCATTAATTAAGTGAATTTTCTGCGCTGAGACTTGACAGGAGCAAGA 60

QY 61 AGACACTCTGGACACCACTATGACAG 87

DB 61 AGACACTCTGGACACCACTATGACAG 87

RESULT 2
AAC55312

ID AAC5312 standard, cDNA, 2818 BP.
 AC AAC5312;
 XX
 XX
 DT 05-FEB-2001 (first entry)
 DE Human activation-induced cytidine deaminase encoding cDNA SEQ ID NO:7.
 XX
 XX Activation-induced cytidine deaminase; AID; cytidine deaminase;
 KM immune related disease; allergy; allergic disease; anti-allergic;
 KM antianaemic; antiasthmatic; ophthalmological; anti-HIV; dermatological;
 KM gene therapy; B cell associated immune system disorder; food allergy;
 KM immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
 KM IGA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis;
 KM drug allergy; allergic rhinitis; Rosen disease; Digeorge disease; AIDS;
 KM ataxia telangiectasia; common variable immunodeficiency disorder;
 KM major histocompatibility class II deficiency disease;
 KM auto immunodeficiency syndrome; IgG subclass selection disorder; ss.
 XX
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 80..676
 FT /tag= a
 FT /product= "activation-induced cytidine deaminase"
 XX
 XX MO200058480-A1.
 XX
 PD 05-OCT-2000.
 PF 28-MAR-2000; 2000WO-JP001918.
 XX
 XX 29-MAR-1999; 99JP-00087192.
 PR 24-JUN-1999; 99JP-00178999.
 PR 27-DEC-1999; 99JP-00371382.
 XX
 PA (NIBS) JAPAN TOBACCO INC.
 PA (HONJ/) HONJO T.
 FI Honjo T, Muramatsu M;
 DR WPI: 2000-611715/58.
 DR P-PsDB; AAB24198.
 XX
 PT Nucleic acid encoding activation induced cytidine deaminase, useful as a
 PT target for drug development for immune-related diseases including
 PT allergies.
 XX
 PS Claim 3, Page 135-139; 174pp; Japanese.
 XX
 CC The present sequence encodes human activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has antiallergic, antianaemic,
 CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
 CC globulinaemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, Digeorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, AIDS (auto
 CC histocompatibility class II deficiency disease, elevated IGE disorder, food
 CC immunodeficiency syndrome), elevated IGE disorder, and IgG subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders
 CC
 XX Sequence 2818 BP, 868 A, 548 C, 626 G, 776 T, 0 U, 0 Other;
 SQ
 Query Match 100.0%; Score 87; DB 3; Length 2818;
 Best Local Similarity 100.0%; Pred. No. 9.5e-37;
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAACCATCTATTAATGAGATGATTTTCTGGCCTGAGACTTGACGAGGAGCAGA 60
 DB 1 AGAGAACCATCTATTAATGAGATGATTTTCTGGCCTGAGACTTGACGAGGAGCAGA 60
 QY 61 AGACACTCTGACACCACTATGACAG 87
 DB 61 AGACACTCTGACACCACTATGACAG 87
 XX
 XX
 DE Human activation-induced cytidine deaminase genomic DNA SEQ ID NO:9.
 XX
 XX Activation-induced cytidine deaminase; AID; cytidine deaminase;
 KM immune related disease; allergy; allergic disease; antiallergic;
 KM antianaemic; antiasthmatic; ophthalmological; anti-HIV; dermatological;
 KM gene therapy; B cell associated immune system disorder; food allergy;
 KM immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
 KM IGA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis;
 KM drug allergy; allergic rhinitis; Rosen disease; Digeorge disease; AIDS;
 KM ataxia telangiectasia; common variable immunodeficiency disorder;
 KM major histocompatibility class II deficiency disease;
 KM auto immunodeficiency syndrome; IgG subclass selection disorder; ds.
 XX
 XX Homo sapiens.
 XX
 PN MO200058480-A1.
 PD 05-OCT-2000.
 PF 28-MAR-2000; 2000WO-JP001918.
 XX
 XX 29-MAR-1999; 99JP-00087192.
 PR 24-JUN-1999; 99JP-00178999.
 PR 27-DEC-1999; 99JP-00371382.
 XX
 PA (NIBS) JAPAN TOBACCO INC.
 PA (HONJ/) HONJO T.
 FI Honjo T, Muramatsu M;
 DR WPI: 2000-611715/58.
 DR 29-MAR-1999; 99JP-00087192.
 PR 24-JUN-1999; 99JP-00178999.
 PR 27-DEC-1999; 99JP-00371382.
 XX
 PT Nucleic acid encoding activation induced cytidine deaminase, useful as a
 PT target for drug development for immune-related diseases including
 PT allergies.
 XX
 PS Claim 17, Page 142-145; 174pp; Japanese.
 XX
 CC The present invention describes an activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has antiallergic, antianaemic,
 CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
 CC globulinaemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, Digeorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, AIDS (major
 CC histocompatibility class II deficiency disease, elevated IGE disorder, food
 CC immunodeficiency syndrome), elevated IGE disorder, and IgG subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders. The present sequence represents a
 CC genomic DNA sequence of human AID
 XX

(e.g., rheumatoid arthritis or cancer) associated with mutant protein or cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g., p53), or selectively treating cells expressing mutant protein or cellular protein isoform in a patient heterozygous for (11). The method is useful for treating a disease e.g., hematopoietic disorder such as T or B cell lymphoma, chronic myeloid leukaemia (CML), APL, ALL, AML, NHL and CML, or a disease characterized by a solid tumour such as papillary thyroid carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and synovial sarcoma. The method is also useful for treating viral infections. This represents the DNA sequence of a chromosome aberration

Sequence 11204 BP; 3305 A; 2273 C; 2373 G; 3253 T; 0 U; 0 Other;

Query Match 100.0%; Score 87; DB 6; Length 11204;
Best Local Similarity 100.0%; Pred. No. 9.8e-37; Mismatches 0; Indels 0; Gaps 0;
Matches 87; Conservative 0;

QY 1 AGAGAACATCATTAATTAAGTGAATTTTCTGCGCTGAGACTTGGAGGAGGCAAGA 60
DB 442 AGAGAACATCATTAATTAAGTGAATTTTCTGCGCTGAGACTTGGAGGAGGCAAGA 501
QY 61 AGACACTCTGACACCACTATGACAG 87
DB 502 AGACACTCTGACACCACTATGACAG 528

RESULT 6
ABX05468
ID ABX05468 standard; cDNA; 1543 BP.

AC ABX05468;

DT 17-JAN-2003 (first entry)

DE Human novel polynucleotide #483.

XX Human; gene; ser; genetic disorder; gene mapping; medical imaging; cancer;
KM neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
KM Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
KM osteoarthritis; periodontal disease; liver fibrosis; viral infection;
KM fungal infection; bacterial infection; autoimmune disease; diabetes;
KM atopic dermatitis.

XX Homo sapiens.

PN WO200274961-A1.

XX 26-SEP-2002.

XX 14-MAR-2002; 2002WO-US005109.

XX 15-MAR-2001; 2001US-00810173.

XX (HYSE-) HYSEO INC.

XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F,
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
PI Wehrman T, Wang J, Wang D, Dermanac RT;

XX WPI; 2003-040556/03.

XX P-PSDB; ABU00390.

XX New isolated polypeptides and polynucleotides, useful for preventing,
PT treating or ameliorating medical conditions, such as cancer,
PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
PT disorders, and infections.

XX Claim 1; SEQ ID NO 483; 235pp; English.

XX The invention relates to human polynucleotides and the polypeptides they
CC encode. The polynucleotides and polypeptides are useful in diagnostics,
CC forensics, gene mapping, medical imaging, identification of mutations
CC responsible for genetic disorders or other traits, assessing biodiversity

CC and producing many other types of data and products dependent on DNA and
CC amino acid sequences. They are also useful for preventing, treating or
CC ameliorating medical conditions, such as cancer, neurodegenerative
CC disorders (e.g., Parkinson's disease, Alzheimer's disease), lymphoid cell
CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
CC periodontal disease, liver fibrosis, infections (e.g., viral, fungal or
CC bacterial) or autoimmune diseases (e.g., diabetes, atopic dermatitis).
CC Sequences ABX04986-ABX0511 represent human polynucleotides of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification but is based on sequence information supplied
CC by the European Patent Office

Sequence 1543 BP; 428 A; 334 C; 352 G; 429 T; 0 U; 0 Other;

Query Match 96.6%; Score 84; DB 7; Length 1543;
Best Local Similarity 100.0%; Pred. No. 3.8e-35; Mismatches 0; Indels 0; Gaps 0;
Matches 84; Conservative 0;

QY 4 GAACATCATTAATTAAGTGAATTTTCTGCGCTGAGACTTGGAGGAGGCAAGA 63
DB 2 GAACATCATTAATTAAGTGAATTTTCTGCGCTGAGACTTGGAGGAGGCAAGA 61
QY 64 CACTCTGACACCACTATGACAG 87
DB 62 CACTCTGACACCACTATGACAG 85

RESULT 7
ABX73287
ID ABX73287 standard; DNA; 2791 BP.

AC ABX73287;

DT 04-DEC-2002 (first entry)

DE DNA encoding human translocation del(12p) protein #2.

XX Chromosome aberration; oncogenic fusion protein; cancer;
KM proliferative disease; cellular protein isoform; heat shock protein 90;
KM HSP-90; rheumatoid arthritis; cancer; hematopoietic disorder;
KM T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
KM acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
KM acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
KM papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
KM rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.

XX Homo sapiens.

PN WO200269900-A2.

XX 12-SEP-2002.

XX 01-MAR-2002; 2002WO-US006518.

XX 01-MAR-2001; 2001US-0272751P.

XX (CONF-) CONFORMA THERAPEUTICS CORP.

XX Fritz LC, Burrows FU;

XX WPI; 2002-698710/75.

XX P-PSDB; ABG95083.

XX Treating genetically-defined disease associated with chromosomal
PT aberrations yielding oncogenic fusion proteins, e.g., cell proliferative
PT diseases, involves administering an inhibitor of heat shock protein 90.

XX Disclosure; Page 246-247; 389pp; English.

XX The invention describes a method of treating genetically-defined disease
CC associated with chromosomal aberrations yielding oncogenic fusion
CC proteins (1), treating cancerous cells containing (1) in a heterogeneous
CC cell population, treating proliferative diseases associated with mutant

CC protein or cellular protein isoforms (II) dependent on heat shock protein
CC (HSP)-90, or selectively treating cells expressing (II) involving
CC administering HSP90-inhibitor. The method is useful for treating
CC genetically-defined disease with chromosomal aberration yielding
CC oncogenic fusion protein, treating cancerous cells containing fusion
CC protein in heterogeneous cell population, treating proliferative disease
CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
CC p53), or selectively treating cells expressing mutant protein or cellular
CC protein isoform in a patient heterozygous for (II). The method is useful
CC for treating a disease e.g. hematopoietic disorder such as T or B cell
CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CML,
CC or a disease characterised by a solid tumour such as papillary thyroid
CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
CC synovial sarcoma. The method is also useful for treating viral
CC infections. This represents the DNA sequence of a chromosome aberration
XX
SQ Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;
Query Match 96.6%; Score 84; DB 6; Length 2791;
Best Local Similarity 100.0%; Pred. No. 3.9e-35;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 GAACCATCATTAATTAAGTAGAGATTTTCTGCGCTGAGACTTGACGGAGGCAAGAGA 63
DB 1 GAACCATCATTAATTAAGTAGAGATTTTCTGCGCTGAGACTTGACGGAGGCAAGAGA 60
QY 64 CACTCTGGACACCACTATGACAG 87
DB 61 CACTCTGGACACCACTATGACAG 84
RESULT 8
ABS73288
ID ABS73288 standard; DNA; 2791 BP.
XX
AC ABS73288;
XX
DT 04-DEC-2002 (first entry)
XX
DE DNA encoding human translocation del(12p) protein #3.
XX
XX Chromosome aberration; oncogenic fusion protein; cancer;
KM proliferative disease; cellular protein isoform; heat shock protein 90;
KM HSP-90; rheumatoid arthritis; cancer; hematopoietic disorder;
KM T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
KM acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
KM acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
KM papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
KM rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200269900-A2.
XX
PD 12-SEP-2002.
XX
PF 01-MAR-2002; 2002WO-US006518.
XX
PR 01-MAR-2001; 2001US-0272751P.
XX
PA (CONF-) CONFORMA THERAPEUTICS CORP.
XX
PI Fritz LC, Burrows FJ;
XX
XX WPI; 2002-698710/75.
DR P-PSDB; ABG95084.
XX
PT Treating genetically-defined disease associated with chromosomal
PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
PT diseases, involves administering an inhibitor of heat shock protein 90.
XX
PS Disclosure; Page 246-249; 389pp; English.

XX The invention describes a method of treating genetically-defined disease
CC associated with chromosomal aberrations yielding oncogenic fusion
CC proteins (II), treating cancerous cells containing (I) in a heterogeneous
CC cell population, treating proliferative diseases associated with mutant
CC protein or cellular protein isoforms (II) dependent on heat shock protein
CC (HSP)-90, or selectively treating cells expressing (II) involving
CC administering HSP90-inhibitor. The method is useful for treating
CC genetically-defined disease with chromosomal aberration yielding
CC oncogenic fusion protein, treating cancerous cells containing fusion
CC protein in heterogeneous cell population, treating proliferative disease
CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
CC p53), or selectively treating cells expressing mutant protein or cellular
CC protein isoform in a patient heterozygous for (II). The method is useful
CC for treating a disease e.g. hematopoietic disorder such as T or B cell
CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CML,
CC or a disease characterised by a solid tumour such as papillary thyroid
CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
CC synovial sarcoma. The method is also useful for treating viral
CC infections. This represents the DNA sequence of a chromosome aberration
XX
SQ Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;
Query Match 96.6%; Score 84; DB 6; Length 2791;
Best Local Similarity 100.0%; Pred. No. 3.9e-35;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 GAACCATCATTAATTAAGTAGAGATTTTCTGCGCTGAGACTTGACGGAGGCAAGAGA 63
DB 1 GAACCATCATTAATTAAGTAGAGATTTTCTGCGCTGAGACTTGACGGAGGCAAGAGA 60
QY 64 CACTCTGGACACCACTATGACAG 87
DB 61 CACTCTGGACACCACTATGACAG 84
RESULT 9
AAC55335
ID AAC55335 standard; DNA; 26 BP.
XX
AC AAC55335;
XX
DT 05-FEB-2001 (first entry)
XX
DE Human activation-induced cytidine deaminase PCR primer SEQ ID NO:31.
XX
XX Activation-induced cytidine deaminase; AID; cytidine deaminase;
KM immune related disease; allergy; allergic disease; anti-allergic;
KM anti-anemic; anti-asthmatic; ophthalmological; anti-HIV; dermatological;
KM gene therapy; B cell associated immune system disorder; food allergy;
KM immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
KM Iga nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis;
KM drug allergy; allergic rhinitis; Rosen disease; DiGeorge disease; AIDS;
KM ataxia telangiectasia; common variable immunodeficiency disorder;
KM major histocompatibility class II deficiency disease; PCR primer;
KM auto immunodeficiency syndrome; Igg sublaas selection disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200058480-A1.
XX
PD 05-OCT-2000.
XX
PF 28-MAR-2000; 2000WO-JP001918.
XX
XX 29-MAR-1999; 99JP-00087192.
PR 24-JUN-1999; 99JP-00178999.
PR 27-DEC-1999; 99JP-00371382.
XX
XX (NIBS) JAPAN TOBACCO INC.
XX (HONJ/) HONJO T.
XX

PI Honjo T, Muramatsu M;
 XX
 DR MPI; 2000-611715/58.
 XX
 PT Nucleic acid encoding activation induced cytidine deaminase, useful as a
 PT target for drug development for immune-related diseases including
 PT allergies.
 XX
 PS Claim 20; Page 161; 174pp; Japanese.
 XX
 CC The present invention describes an activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic,
 CC antiaesthetic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
 CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, Digeorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, MHC (major
 CC histocompatibility class II deficiency disease, AIDS (auto
 CC immunodeficiency syndrome), elevated IGE disorder, and IGE subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders. The present sequence represents a
 CC specifically claimed PCR primer for human AID, which is also used in an
 CC example from the present invention
 XX
 SQ Sequence 26 BP; 7 A; 2 C; 6 G; 11 T; 0 U; 0 Other;
 QY Query Match 29.9%; Score 26; DB 3; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.00049;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 1 CATTAATGAAGTGAAGATTCTCTGG 26
 QY 11 CATTAATGAAGTGAAGATTCTCTGG 36
 DB 1 CATTAATGAAGTGAAGATTCTCTGG 26
 RESULT 10
 ID ADB88945 standard; DNA; 20 BP.
 XX
 AC ADB88945;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Activation-induced cytidine deaminase mutagenesis method primer #7.
 XX
 DE ss; cytosstatic; virucide; antidote; mutation;
 KM activation-induced cytidine deaminase; cell phenotype;
 KM monoclonal antibody; antigen; cross-reactivity; tumour;
 KM germ warfare agent; toxin; V region; primer.
 XX
 OS Homo sapiens.
 XX
 PN WO2003061363-A2.
 XX
 PD 31-JUL-2003.
 XX
 PF 15-JAN-2003; 2003MO-US001149.
 XX
 PR 17-JAN-2002; 2002US-0350269P.
 XX
 XX (YESH) UNIV YESHIVA EINSTEIN COLLEGE.
 XX
 PI Martin A, Scharff MD;
 XX
 DR MPI; 2003-747983/70.
 XX
 PT Induction of mutations in gene expressed in eukaryotic cell useful e.g.
 PT to produce antibodies with greater affinity or specificity for antigens
 PT

PT e.g. human pathogens by expressing an activation-induced cytidine
 PT deaminase gene in cell.
 XX
 PS Example 4; Page 34; 73pp; English.
 XX
 CC The invention relates to the induction of mutations in a gene expressed
 CC in a eukaryotic cell, where the gene is operably linked to a promoter and
 CC within 2 kb of promoter, by expressing a transgenic activation-induced
 CC cytidine deaminase (AID) gene in the cell. The method is useful for
 CC producing mutated genes and proteins, determining the effect of a
 CC mutation on a protein or cell phenotype, and producing genes and proteins
 CC with altered properties. It is especially useful for producing antibodies
 CC with altered affinities or specificities for an antigen, or
 CC increased/decreased cross-reactivity for a second antigen, e.g. to
 CC produce antibodies with greater affinity/specificity useful
 CC therapeutically to treat tumours, viruses such as Ebola and Lhaas Fever
 CC or against germ warfare agents, toxins (e.g. ricin). The method may
 CC similarly be used to alter affinity, specificity or cross-reactivity of a
 CC monoclonal antibody. This sequence represents a PCR primer used to
 CC amplify fragments of the V region of monoclonal antibody genes and the
 CC AID gene for use in the method of the invention to induce mutations in
 CC the sequence.
 XX
 SQ Sequence 20 BP; 7 A; 4 C; 7 G; 2 T; 0 U; 0 Other;
 QY Query Match 23.0%; Score 20; DB 8; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.82;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 1 GAGGCAAGAAAGACACTCTGG 20
 QY 52 GAGGCAAGAAAGACACTCTGG 71
 DB 1 GAGGCAAGAAAGACACTCTGG 20
 Search completed: March 13, 2004, 05:37:24
 Job time : 310 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2004, 05:09:14 ; Search time 60 Seconds
(without alignments)
804.679 Million cell updates/sec

Title: US-09-966-880A-11

Perfect score: 87
Sequence: 1 agagaccaccatcatatga.....ctgacaccactatgacag 87

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 20

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B COMB.seq:*
3: /cgn2_6/prodata/2/ina/5A COMB.seq:*
4: /cgn2_6/prodata/2/ina/5B COMB.seq:*
5: /cgn2_6/prodata/2/ina/PTUS COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|---------------|-------|----------------|--------|----|-------------|
| ----- | | | | | |

No matches found

Search completed: March 13, 2004, 09:17:48
Job time : 61 secs

```

RESULT 2
US-09-966-880A-7
; Sequence 7, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Muratsu, Taseuku
; APPLICANT: Muratsu, Masanichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966, 880A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PasteSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 2818
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (80)...(673)
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)...(79)
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (677)...(2818)
US-09-966-880A-7
Query Match 100.0%; Score 87; DB 9; Length 2818
Best Local Similarity 100.0%; Pctd. No. 1,1e-38;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 12:27:59 ; Search time 465.944 Seconds

(without alignments)
13767.261 Million cell updates/sec

Title: US-09-966-880A-12

Perfect score: 148

Sequence: 1 ccccttgatcgaaccgagga.....ttggtatcttcgacataag 148

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb da:*

2: gb htg:*

3: gb in:*

4: gb ov:*

5: gb ov:*

6: gb pat:*

7: gb ph:*

8: gb pl:*

9: gb pr:*

10: gb ro:*

11: gb sts:*

12: gb sy:*

13: gb un:*

14: gb vi:*

15: em ba:*

16: em fun:*

17: em hum:*

18: em in:*

19: em mu:*

20: em om:*

21: em or:*

22: em ov:*

23: em pat:*

24: em ph:*

25: em pl:*

26: em ro:*

27: em sts:*

28: em un:*

29: em vi:*

30: em htg hum:*

31: em htg inv:*

32: em htg other:*

33: em htg mus:*

34: em htg pln:*

35: em htg rod:*

36: em htg mam:*

37: em htg vrt:*

38: em sy:*

39: em htgo hum:*

40: em htgo mus:*

41: em htgo other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 148 | 100.0 | 148 | 6 BD016837 | BD016837 Novel cyt |
| 2 | 148 | 100.0 | 595 | 9 AF529834 | AF529834 Mus muscu |
| 3 | 148 | 100.0 | 596 | 9 AF529815 | AF529815 Homo sapi |
| 4 | 148 | 100.0 | 596 | 9 AF529816 | AF529816 Homo sapi |
| 5 | 148 | 100.0 | 596 | 9 AF529817 | AF529817 Homo sapi |
| 6 | 148 | 100.0 | 596 | 9 AF529818 | AF529818 Homo sapi |
| 7 | 148 | 100.0 | 596 | 9 AF529819 | AF529819 Homo sapi |
| 8 | 148 | 100.0 | 596 | 9 AF529820 | AF529820 Homo sapi |
| 9 | 148 | 100.0 | 596 | 9 AF529821 | AF529821 Homo sapi |
| 10 | 148 | 100.0 | 596 | 9 AF529822 | AF529822 Homo sapi |
| 11 | 148 | 100.0 | 596 | 9 AF529823 | AF529823 Homo sapi |
| 12 | 148 | 100.0 | 596 | 12 AF529830 | AF529830 Mus muscu |
| 13 | 148 | 100.0 | 596 | 12 AF529831 | AF529831 Mus muscu |
| 14 | 148 | 100.0 | 596 | 12 AF529832 | AF529832 Mus muscu |
| 15 | 148 | 100.0 | 596 | 12 AF529833 | AF529833 Mus muscu |
| 16 | 148 | 100.0 | 597 | 12 AF529839 | AF529839 Mus muscu |
| 17 | 148 | 100.0 | 597 | 12 AF529840 | AF529840 Mus muscu |
| 18 | 148 | 100.0 | 597 | 12 AF529841 | AF529841 Cricetulu |
| 19 | 148 | 100.0 | 597 | 12 AF529842 | AF529842 Cricetulu |
| 20 | 148 | 100.0 | 597 | 12 AF529843 | AF529843 Cricetulu |
| 21 | 148 | 100.0 | 597 | 12 AF529844 | AF529844 Cricetulu |
| 22 | 148 | 100.0 | 597 | 12 AF529845 | AF529845 Cricetulu |
| 23 | 148 | 100.0 | 597 | 12 AF529846 | AF529846 Cricetulu |
| 24 | 148 | 100.0 | 597 | 12 AF529847 | AF529847 Cricetulu |
| 25 | 148 | 100.0 | 597 | 12 AF529848 | AF529848 Cricetulu |
| 26 | 148 | 100.0 | 597 | 12 AF529849 | AF529849 Cricetulu |
| 27 | 148 | 100.0 | 597 | 12 AF529850 | AF529850 Cricetulu |
| 28 | 148 | 100.0 | 597 | 12 AF529851 | AF529851 Cricetulu |
| 29 | 148 | 100.0 | 597 | 12 AF529852 | AF529852 Cricetulu |
| 30 | 148 | 100.0 | 597 | 12 AF529853 | AF529853 Cricetulu |
| 31 | 148 | 100.0 | 597 | 12 AF529854 | AF529854 Cricetulu |
| 32 | 148 | 100.0 | 597 | 12 AF529855 | AF529855 Cricetulu |
| 33 | 148 | 100.0 | 597 | 12 AF529856 | AF529856 Cricetulu |
| 34 | 148 | 100.0 | 597 | 12 AF529857 | AF529857 Cricetulu |
| 35 | 148 | 100.0 | 597 | 12 AF529858 | AF529858 Cricetulu |
| 36 | 148 | 100.0 | 597 | 12 AF529859 | AF529859 Cricetulu |
| 37 | 148 | 100.0 | 597 | 12 AF529860 | AF529860 Cricetulu |
| 38 | 148 | 100.0 | 597 | 12 AF529861 | AF529861 Cricetulu |
| 39 | 148 | 100.0 | 597 | 12 AF529862 | AF529862 Cricetulu |
| 40 | 148 | 100.0 | 597 | 12 AF529863 | AF529863 Cricetulu |
| 41 | 148 | 100.0 | 597 | 12 AF529864 | AF529864 Cricetulu |
| 42 | 146.4 | 98.9 | 292 | 9 AB092577 | AB092577 Homo sapi |
| 43 | 146.4 | 98.9 | 591 | 12 AF529855 | AF529855 Cricetulu |
| 44 | 146.4 | 98.9 | 596 | 9 AF529822 | AF529822 Homo sapi |
| 45 | 146.4 | 98.9 | 596 | 9 AF529823 | AF529823 Homo sapi |

ALIGNMENTS

RESULT 1

LOCUS BD016837

DEFINITION Novel cytidine deaminase.

ACCESSION BD016837

VERSION BD016837.1 GI:22558013

KEYWORDS JP 2001245669-A/10.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 148)

AUTHORS Honjo, T. and Muramatsu, M.

TITLE Novel cytidine deaminase

JOURNAL Patent: JP 2001245669-A 10 11-SEP-2001;

```

COMMENT
OS      JAPAN TOBACCO INC,TASUKU HONJOU
OS      Homo sapiens (human)
PN      JP 2001245669-A/10
PD      11-SEP-2001
PF      28-MAR-2000  JP 2000092981
PI      TASUKU HONJOU,MASAMICHI MORIMATSU
PC      C12N15/09,A61K39/395,A61K39/395,A61P1/00,A61P11/06,A61P13/12,
PC      A61P17/00,
PC      A61P27/02,A61P27/16,A61P37/02,A61P37/08,C07K16/18,C12N1/19, PC
C12N1/21,
PC      C12N5/10,C12N9/78,C12P21/02,C12P21/08// (C12N1/21,C12R1:19), PC
(C12N5/10,C12R1:91),C12N15/00,C12N5/00,(C12N5/00,C12R1:91) CC
FH      Key
      Location/Qualifiers
FEATURES
      source
      1..148
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 148; DB 6; Length 148;
Best Local Similarity 100.0%; Pred. No. 1,56-40;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 CCTCTGATGAGACGGAGAGAACTTCTTACCAATTCAAAAATGCGCTGGGCTAAAGG 60
      |||
DB      1 CCTCTGATGAGAACCGAGAGAACTTCTTACCAATTCAAAAATGCGCTGGGCTAAAGG 60
QY      61 TCGGCGTGAAGACCTACCTGTGCTAGTAGTGAAGAGCGGTGACAGTGTACATCCTTTTC 120
      |||
DB      61 TCGGCGTGAAGACCTACCTGTGCTAGTAGTGAAGAGCGGTGACAGTGTACATCCTTTTC 120
QY      121 ACTGACCTTTGGTTATCTTTCGAATPAG 148
      |||
DB      121 ACTGACCTTTGGTTATCTTTCGAATPAG 148
RESULT 2
AF529834 595 bp mRNA linear SYN 17-SEP-2002
LOCUS AF529834
DEFINITION Mus musculus clone 7 transgenic Homo sapiens AID (AID) mRNA,
complete cds.
ACCESSION AF529834
VERSION AF529834.1 GI:22297255
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Murinae; Mus.
1 (bases 1 to 595)
Martin,A. and Schaffr,M.D.
Somatic hypermutation of the AID transgene in B and non-B cells
Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
2 (bases 1 to 595)
Martin,A. and Schaffr,M.D.
Direct Submision
Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Chann 404, Bronx, NY 10461, USA
FEATURES
      source
      1..595
      /organism="Mus musculus"
      /mol_type="mRNA"
      /db_xref="taxon:10090"
      /clone="7"
      /cell_type="hybridoma PL-5"
      /transgenic
      1..595
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      1..595
      /gene="AID"
      1..>595
CDS
gene
CDS

```

| | | | | | |
|--------|-----------------------|-----------------|--------------------|-----------|-------------|
| ORIGIN | Query Match | 100.0%; | Score 148; | DB 12; | Length 595; |
| | Best Local Similarity | 100.0%; | Pred. No. 1.7e-40; | | |
| | Matches 148; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| | Query Match | 100.0%; | Score 148; | DB 9; | Length 596; |
| | Best Local Similarity | 100.0%; | Pred. No. 1.7e-40; | | |
| | Matches 148; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |

QY 1 CCTCTGATGACCGGAGGAAGTTCTTTTACCAATTCATAAATGTCGCCGCTAGAGG 60
 Db 9 CCTCTGATGACCGGAGGAAGTTCTTTTACCAATTCATAAATGTCGCCGCTAGAGG 68
 QY 61 TCGGCGGAGACCTACCTGCTACGTAAGAGAGCGGTGACAGTCAATCCTTTTC 120
 Db 69 TCGGCGGAGACCTACCTGCTACGTAAGAGAGCGGTGACAGTCAATCCTTTTC 128
 QY 121 ACTGACTTTGGTTATCTTCGCAATAAG 148
 Db 129 ACTGACTTTGGTTATCTTCGCAATAAG 156
 RESULT 4
 AF529816 596 bp mRNA linear PRI 17-SEP-2002
 LOCUS Homo sapiens clone Ramos 2 AID (AID) mRNA, partial cds.
 DEFINITION AF529816 GI:22297219
 VERSION AF529816.1 GI:22297219
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 596)
 Martin, A. and Scharff, M.D.
 Somatic hypermutation of the AID transgene in B and non-B cells
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
 2 (bases 1 to 596)
 Martin, A. and Scharff, M.D.
 Direct Submision
 TITLE Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
 JOURNAL Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA
 FEATURES
 Location/Qualifiers
 1..596
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="Ramos 2"
 1..>596
 /gene="AID"
 1..>596
 /note="Integrated into Burkitt's lymphoma cell line Ramos"
 /codon_start=1
 /product="AID"
 /protein_id="AA95403.1"
 /db_xref="GI:22297220"
 /translation="MDSILMNRKPLVOPNVMAKGRRETYLCVYVKRDSATPSL
 DFGYLRNKGCHVELFLRYSIDMDLPGRCRYVTFTRSPYDCARHVAIDLKGNP
 NLSLRIFLARLYFCEDRKAPBEGRLRLRAGVQIALMTFQDFYCWNTFVENHERTEK
 AMEGLHENSVALSRQLRRLILPLVEVDLDRDARFTLGL"
 ORIGIN
 Query Match 100.0%; Score 148; DB 9; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.7e-40; Indels 0; Gaps 0;
 Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCTCTGATGACCGGAGGAAGTTCTTTTACCAATTCATAAATGTCGCCGCTAGAGG 60
 Db 9 CCTCTGATGACCGGAGGAAGTTCTTTTACCAATTCATAAATGTCGCCGCTAGAGG 68
 QY 61 TCGGCGGAGACCTACCTGCTACGTAAGAGAGCGGTGACAGTCAATCCTTTTC 120
 Db 69 TCGGCGGAGACCTACCTGCTACGTAAGAGAGCGGTGACAGTCAATCCTTTTC 128
 QY 121 ACTGACTTTGGTTATCTTCGCAATAAG 148
 Db 129 ACTGACTTTGGTTATCTTCGCAATAAG 156
 RESULT 5
 AF529817

LOCUS AF529817 596 bp mRNA linear PRI 17-SEP-2002
 DEFINITION Homo sapiens clone Ramos 3 AID (AID) mRNA, partial cds.
 ACCESSION AF529817
 VERSION AF529817.1 GI:22297221
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 596)
 Martin, A. and Scharff, M.D.
 Somatic hypermutation of the AID transgene in B and non-B cells
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
 2 (bases 1 to 596)
 Martin, A. and Scharff, M.D.
 Direct Submision
 TITLE Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
 JOURNAL Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA
 FEATURES
 Location/Qualifiers
 1..596
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="Ramos 3"
 1..>596
 /gene="AID"
 1..>596
 /note="Integrated into Burkitt's lymphoma cell line Ramos"
 /codon_start=1
 /product="AID"
 /protein_id="AA95404.1"
 /db_xref="GI:22297222"
 /translation="MDSILMNRKPLVOPNVMAKGRRETYLCVYVKRDSATPSL
 DFGYLRNKGCHVELFLRYSIDMDLPGRCRYVTFTRSPYDCARHVAIDLKGNP
 NLSLRIFLARLYFCEDRKAPBEGRLRLRAGVQIALMTFQDFYCWNTFVENHERTEK
 AMEGLHENSVALSRQLRRLILPLVEVDLDRDARFTLGL"
 ORIGIN
 Query Match 100.0%; Score 148; DB 9; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.7e-40; Indels 0; Gaps 0;
 Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCTCTGATGACCGGAGGAAGTTCTTTTACCAATTCATAAATGTCGCCGCTAGAGG 60
 Db 9 CCTCTGATGACCGGAGGAAGTTCTTTTACCAATTCATAAATGTCGCCGCTAGAGG 68
 QY 61 TCGGCGGAGACCTACCTGCTACGTAAGAGAGCGGTGACAGTCAATCCTTTTC 120
 Db 69 TCGGCGGAGACCTACCTGCTACGTAAGAGAGCGGTGACAGTCAATCCTTTTC 128
 QY 121 ACTGACTTTGGTTATCTTCGCAATAAG 148
 Db 129 ACTGACTTTGGTTATCTTCGCAATAAG 156
 RESULT 6
 AF529818 596 bp mRNA linear PRI 17-SEP-2002
 LOCUS AF529818
 DEFINITION Homo sapiens clone Ramos 4 truncated AID (AID) mRNA, complete cds.
 ACCESSION AF529818
 VERSION AF529818.1 GI:22297223
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 596)
 Martin, A. and Scharff, M.D.
 Somatic hypermutation of the AID transgene in B and non-B cells
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
 2 (bases 1 to 596)
 Martin, A. and Scharff, M.D.

TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA

FEATURES
source
1..596
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="Ramos 4"
1..596
/gene="AID"
1..570
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="truncated AID"
/protein_id="AA095405.1"
/db_xref="GI:22297224"
/translation="MDSLIMNRKFLYOFKNVWAKGRRETYLCYVVKRDSATSPSL
DFGLRNKNGCHVELFLRYISDMDDPGCYRVTFWTSNSPCYDCARHVDPLRGNP
NLSLRITFARLYFCEDHKAPBEGRLRHRAVGQIALMTPEKDYFCMNTFVENHERTFK
AMEGLHENSVALSRQLRIILPLVEVDL"

ORIGIN
Query Match 100.0%; Score 148; DB 9; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.7e-40; Indels 0; Gaps 0;
Matches 148; Conservative 0; Mismatches 0;

QY 1 CCTCTTGATGACCGGAGGAAGTTCTTTACCAATTCAAAATGTCGCTGGGCTAAGG 60
DB 9 CCTCTTGATGACCGGAGGAAGTTCTTTACCAATTCAAAATGTCGCTGGGCTAAGG 68

QY 61 TCGGCGTGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 69 TCGGCGTGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 128

QY 121 ACTGACTTGGTTATCTTCCCAATAG 148
DB 129 ACTGACTTGGTTATCTTCCCAATAG 156

RESULT 7
AF529819 596 bp mRNA linear PRI 17-SEP-2002
LOCUS Homo sapiens clone Ramos 5 AID (AID) mRNA, partial cds.
ACCESSION AF529819
VERSION AF529819.1 GI:22297225

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
REFERENCE 2 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA

FEATURES
source
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 5"
1..596
/gene="AID"
1..596
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1

/product="AID"
/protein_id="AA095406.1"
/db_xref="GI:22297226"
/translation="MDSLIMNRKFLYOFKNVWAKGRRETYLCYVVKRDSATSPSL
DFGLRNKNGCHVELFLRYISDMDDPGCYRVTFWTSNSPCYDCARHVDPLRGNP
NLSLRITFARLYFCEDHKAPBEGRLRHRAVGQIALMTPEKDYFCMNTFVENHERTFK
AMEGLHENSVALSRQLRIILPLVEVDL"

ORIGIN
Query Match 100.0%; Score 148; DB 9; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.7e-40; Indels 0; Gaps 0;
Matches 148; Conservative 0; Mismatches 0;

QY 1 CCTCTTGATGACCGGAGGAAGTTCTTTACCAATTCAAAATGTCGCTGGGCTAAGG 60
DB 9 CCTCTTGATGACCGGAGGAAGTTCTTTACCAATTCAAAATGTCGCTGGGCTAAGG 68

QY 61 TCGGCGTGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 69 TCGGCGTGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 128

QY 121 ACTGACTTGGTTATCTTCCCAATAG 148
DB 129 ACTGACTTGGTTATCTTCCCAATAG 156

RESULT 8
AF529820 596 bp mRNA linear PRI 17-SEP-2002
LOCUS Homo sapiens clone Ramos 6 AID (AID) mRNA, partial cds.
ACCESSION AF529820
VERSION AF529820.1 GI:22297227

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
REFERENCE 2 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA

FEATURES
source
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 6"
1..596
/gene="AID"
1..596
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="AA095407.1"
/db_xref="GI:22297228"
/translation="MDSLIMNRKFLYOFKNVWAKGRRETYLCYVVKRDSATSPSL
DFGLRNKNGCHVELFLRYISDMDDPGCYRVTFWTSNSPCYDCARHVDPLRGNP
NLSLRITFARLYFCEDHKAPBEGRLRHRAVGQIALMTPEKDYFCMNTFVENHERTFK
AMEGLHENSVALSRQLRIILPLVEVDL"

ORIGIN
Query Match 100.0%; Score 148; DB 9; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.7e-40; Indels 0; Gaps 0;
Matches 148; Conservative 0; Mismatches 0;

QY 1 CCTCTTGATGACCGGAGGAAGTTCTTTACCAATTCAAAATGTCGCTGGGCTAAGG 60
DB 9 CCTCTTGATGACCGGAGGAAGTTCTTTACCAATTCAAAATGTCGCTGGGCTAAGG 68

Db 9 CCTCTTATGATGACCGAGGAAGTTCTTTACCAATTCAAAAATGTCGGCTGGCTAAGG 68
QY 61 TCGGCGTGAGACCTACCTGCTCTAGTAGTAGAGAGGCGTGAACAGTCTCATCTCTTTC 120
Db 69 TCGGCGTGAGACCTACCTGCTCTAGTAGTAGAGAGGCGTGAACAGTCTCATCTCTTTC 128
QY 121 ACTGACCTTTGGTTATCTTCTGCAATTAAG 148
Db 129 ACTGACCTTTGGTTATCTTCTGCAATTAAG 156

RESULT 9
AF529821 596 bp mRNA linear PRI 17-SEP-2002
LOCUS Homo sapiens clone Ramos 7 AID (AID) mRNA, partial cds.
AF529821
VERSION AF529821.1 GI:22297229
KEYWORDS
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
REFERENCE 2 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Ave. Bronx, NY 10461, USA

FEATURES
source
1..596
location/Qualifiers
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 7"
1..>596
/gene="AID"
1..>596
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/product="AID"
/protein_id="AA095408.1"
/db_xref="GI:22297230"
/translation="MDSILNMRKFLYQFNVRNKAQRRETYLCYVVKRDSATSFSLDFGILRNKNGHVELFLRYISDMDPDRGCVRYMTFSPCYDCARHVADELKGNPMLSRLITFARLYFCEDRKAPBEGRLRHARGVOQAIIMTFDYFCMNTFVNHRTFKAMEGHNHNSVRLSRQLRRILPLFYEVDLDRDAFRTTGL"

ORIGIN
Query Match 100.0%; Score 148; DB 9; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.7e-40;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ACCESSION AF529824
VERSION AF529824.1 GI:22297235
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
REFERENCE 2 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Ave. Bronx, NY 10461, USA

FEATURES
source
1..596
location/Qualifiers
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 10"
1..>596
/gene="AID"
1..>596
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/product="AID"
/protein_id="AA095411.1"
/db_xref="GI:22297236"
/translation="MDSILNMRKFLYQFNVRNKAQRRETYLCYVVKRDSATSFSLDFGILRNKNGHVELFLRYISDMDPDRGCVRYMTFSPCYDCARHVADELKGNPMLSRLITFARLYFCEDRKAPBEGRLRHARGVOQAIIMTFDYFCMNTFVNHRTFKAMEGHNHNSVRLSRQLRRILPLFYEVDLDRDAFRTTGL"

ORIGIN
Query Match 100.0%; Score 148; DB 9; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.7e-40;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 CCTCTTATGATGACCGAGGAAGTTCTTTACCAATTCAAAAATGTCGGCTGGCTAAGG 68
QY 61 TCGGCGTGAGACCTACCTGCTCTAGTAGTAGAGAGGCGTGAACAGTCTCATCTCTTTC 120
Db 69 TCGGCGTGAGACCTACCTGCTCTAGTAGTAGAGAGGCGTGAACAGTCTCATCTCTTTC 128
QY 121 ACTGACCTTTGGTTATCTTCTGCAATTAAG 148
Db 129 ACTGACCTTTGGTTATCTTCTGCAATTAAG 156

RESULT 11
AF529826 596 bp mRNA linear PRI 17-SEP-2002
LOCUS Homo sapiens clone Ramos 12 AID (AID) mRNA, partial cds.
AF529826
VERSION AF529826.1 GI:22297239
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
REFERENCE 2 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of

FEATURES
Source
Medicine, 1300 Morris Park Ave, Chanin 404, Bronx, NY 10461, USA
Location/Qualifiers
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 12"
1..>596
/gene="AID"
1..>596
/gene="AID"
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="AAM95414.1"

ORIGIN
Query Match 100.0%; Score 148; DB 9; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.7e-40;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTGATGACCGGAGAGTTCTTTACCAATTGCAAAATGTCGGCTGGCTAAGG 60
DB 9 CCTCTGATGACCGGAGAGTTCTTTACCAATTGCAAAATGTCGGCTGGCTAAGG 68
QY 61 TCGGCGTAGACCTACCTGCTGCTAGTAGAAGGCGTAGACGCTACATCTTTTC 120
DB 69 TCGGCGTAGACCTACCTGCTGCTAGTAGAAGGCGTAGACGCTACATCTTTTC 128
QY 121 ACTGACCTTGGTTATCTTCGCAATAAG 148
DB 129 ACTGACCTTGGTTATCTTCGCAATAAG 156

RESULT 12
AF529827 596 bp mRNA linear PRI 17-SEP-2002
LOCUS Homo sapiens clone Ramos 13 AID (AID) mRNA, partial cds.
DEFINITION AF529827
ACCESSION AF529827.1 GI:22297241
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
2 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Direct Submision
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave, Chanin 404, Bronx, NY 10461, USA

FEATURES
Source
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 13"
1..>596
/gene="AID"
1..>596
/gene="AID"
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="AAM95414.1"

FEATURES
Source
Medicine, 1300 Morris Park Ave, Chanin 404, Bronx, NY 10461, USA
Location/Qualifiers
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 12"
1..>596
/gene="AID"
1..>596
/gene="AID"
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="AAM95414.1"

ORIGIN
Query Match 100.0%; Score 148; DB 9; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.7e-40;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTGATGACCGGAGAGTTCTTTACCAATTGCAAAATGTCGGCTGGCTAAGG 60
DB 9 CCTCTGATGACCGGAGAGTTCTTTACCAATTGCAAAATGTCGGCTGGCTAAGG 68
QY 61 TCGGCGTAGACCTACCTGCTGCTAGTAGAAGGCGTAGACGCTACATCTTTTC 120
DB 69 TCGGCGTAGACCTACCTGCTGCTAGTAGAAGGCGTAGACGCTACATCTTTTC 128
QY 121 ACTGACCTTGGTTATCTTCGCAATAAG 148
DB 129 ACTGACCTTGGTTATCTTCGCAATAAG 156

RESULT 13
AF529830 596 bp mRNA linear SYN 17-SEP-2002
LOCUS Mus musculus clone 3 transgenic Homo sapiens AID (AID) mRNA,
DEFINITION complete cds.
ACCESSION AF529830
VERSION AF529830.1 GI:22297247
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
2 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Direct Submision
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave, Chanin 404, Bronx, NY 10461, USA

FEATURES
Source
1..596
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="3"
/cell_type="hybridoma Pl-5"
/transgenic
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..>596
/gene="AID"
1..>596
/gene="AID"
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="AAM95417.1"
/translation="MDSLMMRRKFLYQFNVMWAKGRRETYLCYVVKRDSATPSL
DFGILRNKNGCHVELLFLRYISDMDDPGCYRVMTFSWSPCYDCAHHVADPLRGNP
NLSLRIFATRLYFCEDRKAPBEGRLRHRAQVQIAIMTFKDYFCWNTFVENHERTFK
AMEGLHNSVRLSRQLRLILPLYEVDDLDAFRTTGL"

| | Matches | 148; | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
|----|---------|---------------------|--|--------------------------|------------|----|--------|----|------|----|
| QY | 1 | CCTCTTGATGAA | CCGGAGGAAGTTCTTTACCAATTC | CAAAAATGTCGGCTGGGCTTAAGG | 60 | | | | | |
| Db | 9 | CCTTTGATGAA | CCGGAGGAAGTTCTTTACCAATTC | CAAAAATGTCGGCTGGGCTTAAGG | 68 | | | | | |
| QY | 61 | TCGGCGTAGA | CCATCCTGTGCTACGTAGTGAAGAGCGGTGACAGTGTACATCCTTTTC | 120 | | | | | | |
| Db | 69 | TCGGCGTAGA | CCATCCTGTGCTACGTAGTGAAGAGCGGTGACAGTGTACATCCTTTTC | 128 | | | | | | |
| QY | 121 | ACTGACCTTGGTTATCTTC | TCCGAATAAG | 148 | | | | | | |
| Db | 129 | ACTGACCTTGGTTATCTTC | TCCGAATAAG | 156 | | | | | | |

```

RESULT 14
AF529831
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

```

```

source
1. 596
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:10090"
/clone="4"
/cell_type="hybridoma Pl-5"
/transgenic
1. 596
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
1. 596
/gene="AID"
1. 596
/gene="AID"
/codon_start=1
/product="AID"
/protein_id="AAM95418.1"
/db_xref="GI:22297250"
/translation="MDSILNMRKRLVQFKNVNMAKGRRETYLCVYVRSDASFSI
DGYLRNKGCHVELLPRIYSDDMDLDGRCRYTWFTSMSPCYDCAHVADPLRGN
NLSNLSIFARLYFCEDRKAPEEGLRRLHRAVQALIKTFKDYFICMNTFVENHETP
AMEGIIHNSVRLSRQFRILLPLEVVDLRARFTLGI"

```

| | | | | | |
|----|--|--------|-------------------|-----------|------------|
| | Query Match | 100.0% | Score 148 | DB 12 | Length 596 |
| | Best Local Similarity | 100.0% | Pred. No. 1.7e-40 | | |
| | Matches 148; Conservative | 0; | Mismatches 0; | Indels 0; | Gaps 0. |
| Qy | 1 CCTCTGATGAACCGAGAAAGTTTCTTACAATTCAAAAATGTCGGTGAGGCTAAGG | 60 | | | |
| Dd | 9 CCTCTGATGAACCGAGAAAGTTTCTTACAATTCAAAAATGTCGGTGAGGCTAAGG | 68 | | | |
| Qy | 61 TTGGGCTAGAACCCTACTGTGCTACGTAGTAGAAGAGCGTGACAGTCAATCCTTTTC | 120 | | | |
| Dd | 69 TTGGGCTAGAACCCTACTGTGCTACGTAGTAGAAGAGCGTGACAGTCAATCCTTTTC | 128 | | | |

```

QY      121 ACTGCACTTGGTTATCTTCGCAATAAG 148
        |||||
Db      129 ACTGCACTTGGTTATCTTCGCAATAAG 156

```

| | |
|------------|--|
| RESULT 15 | |
| AF529837 | |
| LOCUS | AF529837 596 bp mRNA linear SYN 17-SEP-2002 |
| DEFINITION | Mus musculus clone 10 transgenic Homo sapiens AID (AID) mRNA, |
| ACCESSION | AF529837 |
| VERSION | AF529837.1 GI:22257261 |
| KEYWORDS | |
| SOURCE | Mus musculus (house mouse) |
| ORGANISM | Mus musculus |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |
| AUTHORS | 1 (bases 1 to 596) |
| TITLE | Martin,A. and Scharff,M.D. |
| JOURNAL | Somatic hypermutation of the AID transgene in B and non-B cells |
| REFERENCE | Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002) |
| AUTHORS | 2 (bases 1 to 596) |
| TITLE | Martin,A. and Scharff,M.D. |
| JOURNAL | Direct Submision |
| FEATURES | Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Ave. Bronx, NY 10461, USA |
| | Location/Qualifiers |

```

source
1..596
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="10"
/cell_type="hybridoma P1-5"
/transgenic
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..>596
/gene="AID"
1..>596
/gene="AID"
/codon_start=1
/product="AID"
/protein_id="AAM95424.1"
/db_xref="GI:22297262"
/translation="MDSLNNRRKFLYQFNVWAKGRRTYLCTVYKRSATSFSIL
DPEYLRKNGCHVEILFLRYISMDLIDRGRCYVATWTSSPCYDCARHADPFRGNP
NLSIRITFARLYICEDEKKAPEEGIRRIHRAGVQAIATPQDYFYCMTFVNERHETIK
AMEGLNENRSLRSRQIRILLPLYEVDDLDAFRTGL"

```

| Query March | 100.0%; | Score 148; | DB 12; | Length 596; |
|-----------------------|-----------------|---|-----------|-------------|
| Best Local Similarity | 100.0%; | Pred. No. 1,7e-40; | | |
| Matches 148; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 1 | CCTCTTATGAAACCGAGGAAGTTCTTPTACCAATTCAAAATATGCGCTGGCTTAAGG | 60 | |
| | | | | |
| Db | 9 | CCTTTTATGAAACCGAGGAAGTTCTTTCACATTCAAAATGCTCGGCTTAAGG | 68 | |
| | | | | |
| Qy | 61 | TGCGCGTGAACCTACTCTGTGCTACGTAGTGAAGGCGTGAACGTCTACTCTTTTC | 120 | |
| | | | | |
| Db | 69 | TGCGCGTGAACCTACTCTGTGCTACGTAGTGAAGGCGTGAACGTCTACTCTTTTC | 128 | |
| | | | | |
| Qy | 121 | ACTGCACTTTGGTTATCTTCGCAATAAG | 148 | |
| | | | | |
| Db | 129 | ACTGCACTTTGGTTATCTTCGCAATAAG | 156 | |
| | | | | |

Search completed: March 12, 2004, 21:22:30
Job time : 466.944 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 12:26:39 ; Search time 59.2641 Seconds

(without alignments)
10609.014 Million cell updates/sec

Title: US-09-966-880A-12

Perfect score: 148
Sequence: 1 ccccttgatgaacgagga.....ttggtatcttcgcaataag 148

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : N_Geneseq_29Jan04:*

1: Geneseq1980s:*\n2: Geneseq1990s:*\n3: Geneseq2000s:*\n4: Geneseq2001as:*\n5: Geneseq2001bs:*\n6: Geneseq2002as:*\n7: Geneseq2003as:*\n8: Geneseq2003bs:*\n9: Geneseq2003cs:*\n10: Geneseq2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 148 | 100.0 | 148 | 3 AAC55316 | AAC55316 Human act |
| 2 | 148 | 100.0 | 597 | 8 ADB88952 | ADB88952 AID gene |
| 3 | 148 | 100.0 | 1543 | 7 ABX05468 | ABX05468 Human nov |
| 4 | 148 | 100.0 | 2791 | 6 ABS73287 | ABS73287 DNA encod |
| 5 | 148 | 100.0 | 2791 | 6 ABS73288 | ABS73288 DNA encod |
| 6 | 148 | 100.0 | 2818 | 3 AAC55312 | AAC55312 Human act |
| 7 | 148 | 100.0 | 6564 | 3 AAC55314 | AAC55314 Human act |
| 8 | 148 | 100.0 | 11204 | 3 AAC55339 | AAC55339 Human act |
| 9 | 148 | 100.0 | 11204 | 6 ABS73286 | ABS73286 DNA encod |
| 10 | 109.6 | 74.1 | 2440 | 3 AAC55307 | AAC55307 Mouse act |
| 11 | 41.2 | 27.8 | 254 | 6 ABL60598 | ABL60598 Human ova |
| 12 | 41.2 | 27.8 | 1446 | 3 AAA12409 | AAA12409 CDNA encd |
| 13 | 41.2 | 27.8 | 1534 | 2 AAZ20856 | AAZ20856 Polynucle |
| 14 | 41.2 | 27.8 | 1534 | 4 AAS59293 | AAS59293 Human CDN |
| 15 | 41.2 | 27.8 | 1534 | 6 ABA90962 | ABA90962 Human pol |
| 16 | 41.2 | 27.8 | 1717 | 9 ADD18987 | ADD18987 Human dis |
| 17 | 41.2 | 27.8 | 201143 | 6 ABR83568 | ABR83568 Human DNA |
| 18 | 40.8 | 27.6 | 154 | 4 AAI27640 | AAI27640 Probe #17 |
| 19 | 40.8 | 27.6 | 154 | 4 ABA75951 | ABA75951 Human foe |
| 20 | 40.8 | 27.6 | 154 | 4 AAI56609 | AAI56609 Probe #25 |
| 21 | 40.8 | 27.6 | 154 | 4 ABA40515 | ABA40515 Probe #18 |
| 22 | 40.8 | 27.6 | 154 | 4 AAK50628 | AAK50628 Human bon |
| 23 | 40.8 | 27.6 | 154 | 4 AAK24631 | AAK24631 Human bra |

| | | | | | |
|----|------|------|-------|-------------|---------------------|
| 24 | 40.8 | 27.6 | 154 | 4 ABE50231 | ABE50231 Human liv |
| 25 | 40.8 | 27.6 | 154 | 2 ABE24102 | ABE24102 Human gen |
| 26 | 40.8 | 27.6 | 300 | 2 AAZ13088 | AAZ13088 Human gen |
| 27 | 40.8 | 27.6 | 997 | 4 AAI18617 | AAI18617 Probe #85 |
| 28 | 40.8 | 27.6 | 997 | 4 ABA63615 | ABA63615 Human foe |
| 29 | 40.8 | 27.6 | 997 | 4 AAI43731 | AAI43731 Probe #12 |
| 30 | 40.8 | 27.6 | 997 | 4 ABA30812 | ABA30812 Probe #92 |
| 31 | 40.8 | 27.6 | 997 | 4 AAK37857 | AAK37857 Human bon |
| 32 | 40.8 | 27.6 | 997 | 4 AAK12141 | AAK12141 Human bra |
| 33 | 40.8 | 27.6 | 997 | 4 ABA37486 | ABA37486 Human liv |
| 34 | 40.8 | 27.6 | 997 | 6 ABE11852 | ABE11852 Human gen |
| 35 | 40.8 | 27.6 | 1526 | 9 ADD18895 | ADD18895 Human dis |
| 36 | 40.8 | 27.6 | 1638 | 6 ABE061187 | ABE061187 Unidentif |
| 37 | 40.4 | 27.3 | 716 | 2 AAZ15848 | AAZ15848 Human gen |
| 38 | 39.4 | 26.6 | 371 | 8 ACH43025 | ACH43025 Human foe |
| 39 | 35.4 | 23.9 | 1348 | 6 ABE96785 | ABE96785 Gene #328 |
| 40 | 32.4 | 21.9 | 2151 | 6 ABE62572 | ABE62572 CDNA sequ |
| 41 | 32.2 | 21.8 | 823 | 9 ADE25658 | ADE25658 Human CDN |
| 42 | 31.8 | 21.5 | 12222 | 4 AAF54045 | AAF54045 Human alp |
| 43 | 30 | 20.3 | 30 | 3 AAC55322 | AAC55322 Human act |
| 44 | 28.8 | 19.5 | 2529 | 4 ABL12752 | ABL12752 Drosophil |
| 45 | 28.8 | 19.5 | 2747 | 4 ABL13132 | ABL13132 Drosophil |

ALIGNMENTS

| | |
|---|--|
| RESULT 1 | |
| AAC55316 | |
| ID AAC55316 standard; DNA; 148 BP. | |
| AC AAC55316; | |
| XX | |
| DT 05-FEB-2001 (first entry) | |
| DE Human activation-induced cytidine deaminase exon 2 SEQ ID NO:12. | |
| XX | |
| KW Activation-induced cytidine deaminase; AID; cytidine deaminase; | |
| KW immune related disease; allergy; allergic disease; dermatologic; | |
| KW antianemic; antianemic; ophthalmological; anti-HIV; dermatological; | |
| KW gene therapy; B cell associated immune system disorder; food allergy; | |
| KW immunodeficiency disease; immunoglobulin A deficiency disease; asthma; | |
| KW Iga nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis; | |
| KW drug allergy; allergic rhinitis; Rosen disease; DiGeorge disease; AIDS; | |
| KW ataxia telangiectasia; common variable immunodeficiency disorder; | |
| KW major histocompatibility class II deficiency disease; | |
| KW auto immunodeficiency syndrome; Igg subclass selection disorder; ds. | |
| OS Homo sapiens. | |
| XX | |
| XX WO200058480-A1. | |
| XX | |
| PD 05-OCT-2000. | |
| XX | |
| PF 28-MAR-2000; 2000WO-JP001918. | |
| XX | |
| XX 29-MAR-1999; 99JP-00087192. | |
| PR 24-JUN-1999; 99JP-00178999. | |
| PR 27-DEC-1999; 99JP-00371382. | |
| XX | |
| PA (NISR) JAPAN TOBACCO INC. | |
| PA (HONJ) HONJO T. | |
| PI Honjo T, Muramatsu M; | |
| XX | |
| DR WPI; 2000-611715/58. | |
| XX | |
| PT Nucleic acid encoding activation induced cytidine deaminase; useful as a | |
| PT target for drug development for immune-related diseases including | |
| XX allergies. | |
| XX | |
| PS Claim 18; Page 150; 174pp; Japanese. | |
| XX | |

CC The present invention describes an activation-induced cytidine deaminase (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has cytidine activity similar to APOBEC-1. AID has anti-allergic, antianemic, CC antisthmatic, ophthalmological, anti-HIV and dermatological activities, CC and can be used in gene therapy. AID polynucleotides are useful in CC methods for identifying drugs for the treatment of B cell associated CC immune system disorders, immunodeficiency diseases and allergies, such as CC immunoglobulin A (IgA) deficiency disease, IgA nephritis, gamma- CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy, CC drug allergy, allergic rhinitis, Rosen disease, Digorge disease, ataxia CC telangiectasia, common variable immunodeficiency disorder, MHC (major CC histocompatibility class II deficiency disease, AIDS (auto CC immunodeficiency syndrome), elevated IgE disorder, and IgG subclass CC selection disorder. The DNA sequences encoding AID may be used for gene CC therapy and the antibodies to the AID protein may be used for diagnosis CC and treatment of these disorders. The present sequence represents the CC exon 2 genomic DNA sequence of human AID

XX Sequence 148 BP, 35 A; 33 C; 38 G; 42 T; 0 U; 0 Other;

SQ Query Match 100.0%; Score 148; DB 3; Length 148;

Best Local Similarity 100.0%; Pred. No. 1.7e-43; Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTGATGACCGAGGAGGAGTTCTTTACCAATTCAAAATGTCGCTGGGCTAAGG 60

Db 1 CCTCTGATGACCGAGGAGGAGTTCTTTACCAATTCAAAATGTCGCTGGGCTAAGG 60

QY 61 TCGGCGTGAGACCTACCTGCTGCTACGTAAGAGGCGTGAAGTGTACATCCTTTTC 120

Db 61 TCGGCGTGAGACCTACCTGCTGCTACGTAAGAGGCGTGAAGTGTACATCCTTTTC 120

QY 121 ACTGACCTTGTGTATCTTGGCAATAAG 148

Db 121 ACTGACCTTGTGTATCTTGGCAATAAG 148

RESULT 2

ADB88952 ADB88952 standard, DNA, 597 BP.

XX ADB88952;

DT 04-DEC-2003 (first entry)

DE AID gene as substrate used in AID mutagenic method.

XX ds; cytostatic; virucide; antidote; mutation;

KM activation-induced cytidine deaminase; cell phenotype;

KM monoclonal antibody; antigen; cross-reactivity; tumour;

KM germ warfare agent; toxin; V region.

XX Homo sapiens.

OS

XX

XX

XX

XX

XX

XX

XX

Location/Qualifiers

replace(9,A)

/tag= a /note= "mutation in PI-5 hybridoma cell"

replace(12,A)

/tag= b /note= "mutation in Ramos cell"

replace(14,C)

/tag= c /note= "mutation in CHO cell"

replace(61,T)

/tag= d /note= "mutation in PI-5 hybridoma cell"

replace(66,A)

/tag= e /note= "mutation in PI-5 hybridoma cell"

replace(73,T)

/tag= f /note= "mutation in CHO cell"

/note= "mutation in PI-5 hybridoma cell"

/tag= g /note= "mutation in Ramos cell"

replace(420,T)

/tag= h /note= "mutation in PI-5 hybridoma cell"

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

replace(79,G)

/tag= g /note= "mutation in Ramos cell"

replace(85,G)

/tag= h /note= "mutation in PI-5 hybridoma cell"

replace(112,C)

/tag= i /note= "mutation in Ramos cell"

replace(118,G)

/tag= j /note= "mutation in PI-5 hybridoma cell"

replace(156,T)

/tag= k /note= "mutation in PI-5 hybridoma cell"

replace(161,A)

/tag= l /note= "mutation in PI-5 hybridoma cell"

replace(165,T,G)

/tag= m /note= "mutation in Ramos cell, T mutation in CHO cell"

replace(178,T)

/tag= n /note= "mutation in Ramos cell"

replace(206,G)

/tag= o /note= "mutation in CHO cell"

replace(209,C)

/tag= p /note= "mutation in Ramos cell"

replace(218,T)

/tag= q /note= "mutation in Ramos cell"

replace(250,C)

/tag= r /note= "mutation in Ramos cell"

replace(261,G)

/tag= s /note= "mutation in Ramos cell"

replace(289,A)

/tag= t /note= "mutation in Ramos cell"

replace(303,T)

/tag= u /note= "mutation in Ramos cell"

replace(329,T)

/tag= v /note= "mutation in Ramos cell"

replace(350,G,T)

/tag= w /note= "T mutation in Ramos cell, G mutation in CHO cell"

replace(356,A)

/tag= x /note= "A mutation in Ramos cell, A mutation in CHO cell"

replace(357,G)

/tag= y /note= "mutation in Ramos cell"

replace(367,T, A)

/tag= z /note= "mutation in PI-5 hybridoma cell"

replace(375,A)

/tag= aa /note= "mutation in CHO cell"

replace(402,A)

/tag= ab /note= "mutation in PI-5 hybridoma cell"

replace(410,T)

/tag= ac /note= "mutation in PI-5 hybridoma cell"

replace(418,T)

/tag= ad /note= "mutation in PI-5 hybridoma cell"

replace(420,T)

/tag= ae /note= "mutation in PI-5 hybridoma cell"

replace(420,T)

/tag= af /note= "mutation in PI-5 hybridoma cell"

replace(420,T)

/tag= ag /note= "mutation in PI-5 hybridoma cell"

replace(420,T)

/tag= ah /note= "mutation in PI-5 hybridoma cell"

replace(420,T)

/tag= ai /note= "mutation in PI-5 hybridoma cell"

replace(420,T)

/tag= aj /note= "mutation in PI-5 hybridoma cell"

replace(420,T)

/tag= ak /note= "mutation in PI-5 hybridoma cell"

FT mutation /tag= ae
/note= "mutation in PI-5 hybridoma cell"
FT replace(421,A)
FT /tag= af
/note= "mutation in CHO cell"
FT replace(430,G)
FT /tag= ag
/note= "mutation in Ramos cell"
FT replace(440,A)
FT /tag= ah
/note= "mutation in CHO cell"
FT replace(449,T)
FT /tag= ai
/note= "mutation in PI-5 hybridoma cell"
FT replace(495,A)
FT /tag= aj
/note= "mutation in CHO cell"
FT replace(517,C)
FT /tag= ak
/note= "mutation in Ramos cell"
FT replace(526,T, A)
FT /tag= al
/note= "T mutation in PI-5 hybridoma cell, A mutation in Ramos cell"
FT replace(533,A)
FT /tag= am
/note= "mutation in PI-5 hybridoma cell"
FT replace(538,A)
FT /tag= an
/note= "mutation in CHO cell"
FT replace(549,A)
FT /tag= ao
/note= "mutation in CHO cell"
FT replace(568,T)
FT /tag= ap
/note= "mutation in Ramos cell"
FT replace(570,T)
FT /tag= aq
/note= "mutation in CHO cell"
FT /note= "mutation in CHO cell"
PN WO2003061363-A2.
PD 31-JUL-2003.
PX 15-JAN-2003; 2003WO-US001149.
PX 17-JAN-2002; 2002US-0350269P.
PX (YESH) UNIV YESHIVA EINSTEIN COLLEGE.
PX Martin A, Scharif MD;
XX WPI; 2003-747983/70.
XX
XX Induction of mutations in gene expressed in eukaryotic cell useful e.g.
PT to produce antibodies with greater affinity or specificity for antigens
PT e.g. human pathogens by expressing an activation-induced cytidine
PT deaminase gene in cell.
XX
XX Example 4; Fig 6; 73pp; English.
XX
XX The invention relates to the induction of mutations in a gene expressed
CC in a eukaryotic cell, where the gene is operably linked to a promoter and
CC within 2 kb of promoter, by expressing a transgenic activation-induced
CC cytidine deaminase (AID) gene in the cell. The method is useful for
CC producing mutated genes and proteins, determining the effect of a
CC mutation on a protein or cell phenotype, and producing genes and proteins
CC with altered properties. It is especially useful for producing antibodies
CC with altered affinities or specificities for an antigen, or
CC increased/decreased cross-reactivity for a second antigen, e.g. to
CC produce antibodies with greater affinity/specificity useful
CC therapeutically to treat tumours, viruses such as Ebola and Lassa Fever
CC or against germ warfare agents, toxins (e.g. ricin). The method may

CC similarly be used to alter affinity, specificity or cross-reactivity of a
CC monoclonal antibody. This sequence represents the AID gene used in the
CC method of the invention to induce mutations in the sequence.
XX
SQ Sequence 597 BP; 128 A; 164 C; 155 G; 150 T; 0 U; 0 Other;
Query Match 100.0%; Score 148; DB 8; Length 597;
Best Local Similarity 100.0%; Pred. No. 2.9e-43;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTCTGATGAAACCGGAGAAAGTTCTTTACCAATTCAAATATGTCGGCTGGGCTAAAGG 60
DB 9 CCTCTGATGAAACCGGAGAAAGTTCTTTACCAATTCAAATATGTCGGCTGGGCTAAAGG 68
QY 61 TCGGCGTGAGACCTACCTGTGCTACGTAGTGAAGAGCGGTGACAGTGTACATCCTTTTC 120
DB 69 TCGGCGTGAGACCTACCTGTGCTACGTAGTGAAGAGCGGTGACAGTGTACATCCTTTTC 128
QY 121 ACTGACCTTTGGTTATCTTCGCAATAG 148
DB 129 ACTGACCTTTGGTTATCTTCGCAATAG 156
RESULT 3
ABX05468
ID ABX05468 standard; cDNA; 1543 BP.
XX
AC ABX05468;
XX
DT 17-JAN-2003 (first entry)
XX
DE Human novel polynucleotide #483.
XX
XX Human; gene; ss; genetic disorder; gene mapping; medical imaging; cancer;
KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;
KW fungal infection; bacterial infection; autoimmune disease; diabetes;
KW atopic dermatitis.
XX
XX Homo sapiens.
XX
XX WO200274961-A1.
PN
PD 26-SEP-2002.
PX
PX 14-MAR-2002; 2002WO-US005109.
PX
PX 15-MAR-2001; 2001US-00810173.
PX
PX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QH, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehman T, Wang J, Wang D, Drmanac RT;
XX
XX WPI; 2003-040556/03.
DR
DR P-PSDB; ABU00390.
XX
XX New isolated polypeptides and polynucleotides, useful for preventing,
PT treating or ameliorating medical conditions, such as cancer,
PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
PT disorders, and infections.
XX
PS Claim 1; SEQ ID NO 483; 235pp; English.
XX
XX The invention relates to human polynucleotides and the polypeptides they
CC encode. The polynucleotides and polypeptides are useful in diagnostics,
CC for forensic, gene mapping, medical imaging, identification of mutations
CC responsible for genetic disorders or other traits, assessing biodiversity
CC and producing many other types of data and products dependent on DNA and
CC amino acid sequences. They are also useful for preventing, treating or
CC ameliorating medical conditions, such as cancer, neurodegenerative

CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
CC Sequences ABX04986-ABX05511 represent human polynucleotides of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification but is based on sequence information supplied
CC by the European Patent Office

XX
SQ Sequence 1543 BP; 428 A; 334 C; 352 G; 429 T; 0 U; 0 Other;

Query Match 100.0%; Score 148; DB 7; Length 1543;
Best Local Similarity 100.0%; Pred. No. 4e-43;

Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTGATGAGACCGGAGGAAGTTCTTACCAATTCAAAAATGTCGCTGAAGG 60

DB 86 CCTCTGATGAGACCGGAGGAAGTTCTTACCAATTCAAAAATGTCGCTGAAGG 145

QY 61 TCGGCGTAGACCTACCTGCTACGTAAGAGGCGTGAAGTCTACATCTTTC 120

DB 146 TCGGCGTAGACCTACCTGCTACGTAAGAGGCGTGAAGTCTACATCTTTC 205

QY 121 ACTGACTTGTGTTATCTTGCATTAAG 148

DB 206 ACTGACTTGTGTTATCTTGCATTAAG 233

RESULT 4

AB573287
ID AB573287 standard; DNA; 2791 BP.

AC AB573287;

DT 04-DEC-2002 (first entry)

XX DNA encoding human translocation del(12p) protein #2.

XX Chromosome aberration; oncogenic fusion protein; cancer;
XX proliferative disease; cellular protein isoform; heat shock protein 90;
XX HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
XX T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
XX acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
XX acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
XX papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
XX rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.

XX Homo sapiens.

XX WO200269900-A2.

XX 12-SEP-2002.

XX 01-MAR-2002; 2002WO-US006518.

XX 01-MAR-2001; 2001US-0272751P.

XX (CONF-) CONFORMA THERAPEUTICS CORP.

XX Fritz LC, Burrows FU;

XX WPI; 2002-698710/75.

XX P-PDB; ABG95083.

XX Treating genetically-defined disease associated with chromosomal
XX aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
XX diseases, involves administering an inhibitor of heat shock protein 90.

XX Disclosure; Page 246-247; 389pp; English.

XX The invention describes a method of treating genetically-defined disease
XX associated with chromosomal aberrations yielding oncogenic fusion
XX proteins (I), treating cancerous cells containing (I) in a heterogeneous

CC cell population, treating proliferative diseases associated with mutant
CC protein or cellular protein isoforms (II) dependent on heat shock protein
CC (HSP)-90, or selectively treating cells expressing (II) involving
CC administering HSP90-inhibitor. The method is useful for treating
CC genetically-defined disease with chromosomal aberration yielding
CC oncogenic fusion protein, treating cancerous cells containing fusion
CC protein in heterogeneous cell population, treating proliferative disease
CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
CC p53), or selectively treating cells expressing mutant protein or cellular
CC protein isoform in a patient heterozygous for (II). The method is useful
CC for treating a disease e.g. haematopoietic disorder such as T or B cell
CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CMML,
CC or a disease characterised by a solid tumour such as papillary thyroid
CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
CC synovial sarcoma. The method is also useful for treating viral
CC infections. This represents the DNA sequence of a chromosome aberration

XX
SQ Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;

Query Match 100.0%; Score 148; DB 6; Length 2791;

Best Local Similarity 100.0%; Pred. No. 4.9e-43; Indels 0; Gaps 0;

Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTGATGAGACCGGAGGAAGTTCTTACCAATTCAAAAATGTCGCTGAAGG 60

DB 85 CCTCTGATGAGACCGGAGGAAGTTCTTACCAATTCAAAAATGTCGCTGAAGG 144

QY 61 TCGGCGTAGACCTACCTGCTACGTAAGAGGCGTGAAGTCTACATCTTTC 120

DB 145 TCGGCGTAGACCTACCTGCTACGTAAGAGGCGTGAAGTCTACATCTTTC 204

QY 121 ACTGACTTGTGTTATCTTGCATTAAG 148

DB 205 ACTGACTTGTGTTATCTTGCATTAAG 232

RESULT 5

AB573288
ID AB573288 standard; DNA; 2791 BP.

AC AB573288;

DT 04-DEC-2002 (first entry)

XX DNA encoding human translocation del(12p) protein #3.

XX Chromosome aberration; oncogenic fusion protein; cancer;
XX proliferative disease; cellular protein isoform; heat shock protein 90;
XX HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
XX T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
XX acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
XX acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
XX papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
XX rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.

XX Homo sapiens.

XX WO200269900-A2.

XX 12-SEP-2002.

XX 01-MAR-2002; 2002WO-US006518.

XX 01-MAR-2001; 2001US-0272751P.

XX (CONF-) CONFORMA THERAPEUTICS CORP.

XX Fritz LC, Burrows FU;

XX WPI; 2002-698710/75.

XX P-PDB; ABG95084.

PT Treating genetically-defined disease associated with chromosomal
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
 PT diseases, involves administering an inhibitor of heat shock protein 90.
 XX
 PS Disclosure, Page 248-249; 389pp; English.
 XX
 CC The invention describes a method of treating genetically-defined disease
 CC associated with chromosomal aberrations yielding oncogenic fusion
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous
 CC cell population, treating proliferative diseases associated with mutant
 CC protein or cellular protein isoforms (II) dependent on heat shock protein
 CC (HSP)-90, or selectively treating cells expressing (II) involving
 CC administering HSP90-inhibitor. The method is useful for treating
 CC genetically-defined disease with chromosomal aberration yielding
 CC oncogenic fusion protein, treating cancerous cells containing fusion
 CC protein in heterogeneous cell population, treating proliferative disease
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
 CC p53), or selectively treating cells expressing mutant protein or cellular
 CC protein isoform in a patient heterozygous for (II). The method is useful
 CC for treating a disease e.g. hematopoietic disorder such as T or B cell
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CMML,
 CC or a disease characterised by a solid tumour such as papillary thyroid
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
 CC synovial sarcoma. The method is also useful for treating viral
 CC infections. This represents the DNA sequence of a chromosome aberration
 CC
 XX
 SQ Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;
 Query Match 100.0%; Score 148; DB 6; Length 2791;
 Best Local Similarity 100.0%; Pred. No. 4.9e-43;
 Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCTTTGATGACCGGAGAGATTCTTTACCAATTCAAAATGTCGCTGGCTAAGG 60
 DB CCTTTGATGACCGGAGAGATTCTTTACCAATTCAAAATGTCGCTGGCTAAGG 144
 QY 61 TCGGCGTAGACCTACCTGCTACGTAGTAAGAGGCGTAGACGTCTCATCTCTTTC 120
 DB TCGGCGTAGACCTACCTGCTACGTAGTAAGAGGCGTAGACGTCTCATCTCTTTC 204
 QY 121 ACTGACCTTGGTTATCTTCGCAATTAAG 148
 DB ACTGACCTTGGTTATCTTCGCAATTAAG 232
 RESULT 6
 ID AAC55312 standard; cDNA; 2818 BP.
 XX
 AC AAC55312;
 XX
 DT 05-FEB-2001 (first entry)
 XX
 DE Human activation-induced cytidine deaminase encoding cDNA SEQ ID NO:7.
 XX
 KM Activation-induced cytidine deaminase; AID; cytidine deaminase;
 KM immune related disease; allergy; allergic disease; antiallergic;
 KM antianemic; antianasthmatic; ophthalmological; anti-HIV; dermatological;
 KM gene therapy; B cell associated immune system disorder; food allergy;
 KM immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
 KM IGA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis;
 KM drug allergy; allergic rhinitis; Rosen disease; DiGeorge disease; AIDS;
 KM ataxia telangiectasia; common variable immunodeficiency disorder;
 KM major histocompatibility class II deficiency disease;
 KM auto immunodeficiency syndrome; Igg subclass selection disorder; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 80..676
 FT /tag= a
 FT /product= "activation-induced cytidine deaminase"

XX
 XX WO200058480-A1.
 XX
 XX PD 05-OCT-2000.
 XX
 XX
 PF 28-MAR-2000; 2000MO-UP001918.
 XX
 XX
 PR 29-MAR-1999; 99JP-00087192.
 XX
 PR 24-JUN-1999; 99JP-00178999.
 XX
 PR 27-DEC-1999; 99JP-00371382.
 XX
 PA (NISE) JAPAN TOBACCO INC.
 PA (HONJ/) HONJO T.
 PI Honjo T, Muramatsu M;
 XX
 DR MPI: 2000-611715/58.
 DR P-PSDB: AAB24198.
 XX
 XX Nucleic acid encoding activation induced cytidine deaminase, useful as a
 PT target for drug development for immune-related diseases including
 PT allergies.
 XX
 XX Claim 3; Page 135-139; 174pp; Japanese.
 XX
 CC The present sequence encodes human activation-induced cytidine deaminase
 CC (Aid). Aid structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. Aid has antiallergic, antianemic,
 CC antianasthmatic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. Aid polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (Iga) deficiency disease, IGA nephritis, gamma-
 CC globulinaemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, DiGeorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, MHC (major
 CC histocompatibility class II deficiency disease, AIDS (auto
 CC immunodeficiency syndrome), elevated Igg disorder, and Igg subclass
 CC selection disorder. The DNA sequences encoding Aid may be used for gene
 CC therapy and the antibodies to the Aid protein may be used for diagnosis
 CC and treatment of these disorders
 CC
 XX
 SQ Sequence 2818 BP; 868 A; 548 C; 626 G; 776 T; 0 U; 0 Other;
 Query Match 100.0%; Score 148; DB 3; Length 2818;
 Best Local Similarity 100.0%; Pred. No. 4.9e-43;
 Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCTTTGATGACCGGAGAGATTCTTTACCAATTCAAAATGTCGCTGGCTAAGG 60
 DB CCTTTGATGACCGGAGAGATTCTTTACCAATTCAAAATGTCGCTGGCTAAGG 147
 QY 61 TCGGCGTAGACCTACCTGCTACGTAGTAAGAGGCGTAGACGTCTCATCTCTTTC 120
 DB TCGGCGTAGACCTACCTGCTACGTAGTAAGAGGCGTAGACGTCTCATCTCTTTC 207
 QY 121 ACTGACCTTGGTTATCTTCGCAATTAAG 148
 DB ACTGACCTTGGTTATCTTCGCAATTAAG 235
 RESULT 7
 ID AAC55314 standard; DNA; 6564 BP.
 XX
 AC AAC55314;
 XX
 DT 05-FEB-2001 (first entry)
 XX
 DE Human activation-induced cytidine deaminase genomic DNA SEQ ID NO:10.
 XX
 KM Activation-induced cytidine deaminase; AID; cytidine deaminase;
 KM immune related disease; allergy; allergic disease; antiallergic;

KM antianaemic; antiasthmatic; ophthalmological; anti-HIV; dermatological;
 KM gene therapy; B cell associated immune system disorder; food allergy;
 KM immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
 KM IGA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis;
 KM drug allergy; allergic rhinitis; Rosen disease; Digorge disease; AIDS;
 KM ataxia telangiectasia; common variable immunodeficiency disease;
 KM major histocompatibility class II deficiency disease;
 KM auto immunodeficiency syndrome; Igg subclass selection disorder; ds.
 OS Homo sapiens.
 XX WO200058480-A1.
 PN 05-OCT-2000.
 PD 28-MAR-2000; 2000WO-JP001918.
 PF 29-MAR-1999; 99JP-00087192.
 PR 24-JUN-1999; 99JP-00178999.
 PR 27-DEC-1999; 99JP-00371382.
 XX (NLSB) JAPAN TOBACCO INC.
 PA (HONJ/) HONJO T.
 PI Honjo T, Muramatsu M;
 XX WPI; 2000-611715/58.
 DR Nucleic acid encoding activation induced cytidine deaminase, useful as a
 PT target for drug development for immune-related diseases including
 PT allergies.
 XX Claim 17; Page 145-150; 174pp; Japanese.
 PS The present invention describes an activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has antiasthmatic, antianaemic,
 CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
 CC globulinaemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, Digorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, MHC (major
 CC histocompatibility class) class II deficiency disease, AIDS (auto
 CC immunodeficiency syndrome), elevated Igg disorder, and Igg subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders. The present sequence represents a
 CC genomic DNA sequence of human AID
 CC XX
 SQ Sequence 6564 BP; 1909 A; 1358 C; 1383 G; 1914 T; 0 U; 0 Other;
 Query Match 100.0%; Score 148; DB 3; Length 6564;
 Best Local Similarity 100.0%; Pred. No. 6.7e-43;
 Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCTCTTGATGAACCGAGAGATTCTTACCAATCAAAATGTCGCTGGCTAAGG 60
 DB 1065 CCTCTTGATGAACCGAGAGATTCTTACCAATCAAAATGTCGCTGGCTAAGG 1124
 QY 61 TCGGCGTAGACCTACCTGTCTACGTAGAGAGCGGTACAGTCTACATCTTTTC 120
 DB 1125 TCGGCGTAGACCTACCTGTCTACGTAGAGAGCGGTACAGTCTACATCTTTTC 1184
 QY 121 ACTGACCTTGGTATCTTCCCAATAG 148
 DB 1185 ACTGACCTTGGTATCTTCCCAATAG 1212

ID AAC55339 standard; DNA; 11204 BP.
 XX AAC55339;
 AC 05-FEB-2001 (first entry)
 XX Human activation-induced cytidine deaminase genomic DNA SEQ ID NO:35.
 DE Activation-induced cytidine deaminase; AID; cytidine deaminase;
 KM immune related disease; allergy; allergic disease; antiasthmatic;
 KM antianaemic; antiasthmatic; ophthalmological; anti-HIV; dermatological;
 KM gene therapy; B cell associated immune system disorder; food allergy;
 KM immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
 KM IGA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis;
 KM drug allergy; allergic rhinitis; Rosen disease; Digorge disease; AIDS;
 KM ataxia telangiectasia; common variable immunodeficiency disorder;
 KM major histocompatibility class II deficiency disease;
 KM auto immunodeficiency syndrome; Igg subclass selection disorder; ds.
 OS Homo sapiens.
 XX WO200058480-A1.
 PN 05-OCT-2000.
 PD 28-MAR-2000; 2000WO-JP001918.
 PF 29-MAR-1999; 99JP-00087192.
 PR 24-JUN-1999; 99JP-00178999.
 PR 27-DEC-1999; 99JP-00371382.
 XX (NLSB) JAPAN TOBACCO INC.
 PA (HONJ/) HONJO T.
 PI Honjo T, Muramatsu M;
 XX WPI; 2000-611715/58.
 DR Nucleic acid encoding activation induced cytidine deaminase, useful as a
 PT target for drug development for immune-related diseases including
 PT allergies.
 XX Claim 17; Page 163-170; 174pp; Japanese.
 PS The present invention describes an activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has antiasthmatic, antianaemic,
 CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
 CC globulinaemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, Digorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, MHC (major
 CC histocompatibility class) class II deficiency disease, AIDS (auto
 CC immunodeficiency syndrome), elevated Igg disorder, and Igg subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders. The present sequence represents a
 CC genomic DNA sequence of human AID
 CC XX
 SQ Sequence 11204 BP; 3305 A; 2273 C; 2373 G; 3253 T; 0 U; 0 Other;
 Query Match 100.0%; Score 148; DB 3; Length 11204;
 Best Local Similarity 100.0%; Pred. No. 8.1e-43;
 Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCTCTTGATGAACCGAGAGATTCTTACCAATCAAAATGTCGCTGGCTAAGG 60
 DB 6280 CCTCTTGATGAACCGAGAGATTCTTACCAATCAAAATGTCGCTGGCTAAGG 6339
 QY 61 TCGGCGTAGACCTACCTGTCTACGTAGAGAGCGGTACAGTCTACATCTTTTC 120

Db 6340 TCGGCGTGAACCTTGTCTGCTAGTGAAGAGCGTGACAGTCTCATCTTTTC 6399
 QY 121 ACTGACCTTGGTTATCTTGGCATTAAG 148
 Db 6400 ACTGACCTTGGTTATCTTGGCATTAAG 6427

RESULT 9
 ABS73286
 ID ABS73286 standard; DNA; 11204 BP.

XX ABS73286;
 AC
 XX 04-DEC-2002 (first entry)
 DE DNA encoding human translocation del(12p) protein #1.

XX Chromosome aberration; oncogenic fusion protein; cancer;
 KW proliferative disease; cellular fusion isoform; heat shock protein 90;
 KW HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
 KW T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
 KW acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
 KW acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
 KW papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
 KW rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.

XX Homo sapiens.
 OS
 XX WO200269900-A2.

XX 12-SEP-2002.
 XX 01-MAR-2002; 2002WO-US006518.
 XX 01-MAR-2001; 2001US-0272751P.

XX (CONF-) CONFORMA THERAPEUTICS CORP.

XX Fultz LC, Burrows FU;
 XX WPI: 2002-696710/75.
 DR P-PSDB; ABG95082.

XX Treating genetically-defined disease associated with chromosomal
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
 PT diseases, involves administering an inhibitor of heat shock protein 90.

XX Disclosure; Page 242-245; 389pp; English.

XX The invention describes a method of treating genetically-defined disease
 CC associated with chromosomal aberrations yielding oncogenic fusion
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous
 CC cell population, treating proliferative diseases associated with mutant
 CC protein or cellular protein isoforms (II) dependent on heat shock protein
 CC (HSP)-90, or selectively treating cells expressing (II) involving
 CC administering HSP90-inhibitor. The method is useful for treating
 CC genetically-defined disease with chromosomal aberration yielding
 CC oncogenic fusion protein, treating cancerous cells containing fusion
 CC protein in heterogeneous cell population, treating proliferative disease
 CC (e.g. Rheumatoid arthritis or cancer) associated with mutant protein or
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
 CC P53), or selectively treating cells expressing mutant protein or cellular
 CC protein isoform in a patient heterozygous for (II). The method is useful
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, AML, NHL and CMML,
 CC or a disease characterised by a solid tumour such as papillary thyroid
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
 CC synovial sarcoma. The method is also useful for treating viral
 CC infections. This represents the DNA sequence of a chromosome aberration
 XX Sequence 11204 BP; 3305 A; 2273 C; 2373 G; 3253 T; 0 U; 0 Other;

Query Match 100.0%; Score 148; DB 6; Length 11204;
 Best Local Similarity 100.0%; Pred. No. 8.1e-43;
 Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTTGATGAACCGGAGAAATTTCTTACCAATTAATAATGCCGCTGGCTAAGG 60
 Db 6280 CCTCTTGATGAACCGGAGAAATTTCTTACCAATTAATAATGCCGCTGGCTAAGG 6339
 QY 61 TCGGCGTGAACCTTGTCTGCTAGTGAAGAGCGTGACAGTGTACATCTTTTC 120
 Db 6340 TCGGCGTGAACCTTGTCTGCTAGTGAAGAGCGTGACAGTGTACATCTTTTC 6399

QY 121 ACTGACCTTGGTTATCTTGGCATTAAG 148
 Db 6400 ACTGACCTTGGTTATCTTGGCATTAAG 6427

RESULT 10
 AAC5307
 ID AAC5307 standard; cDNA; 2440 BP.

XX AAC5307;
 AC
 XX 05-FEB-2001 (first entry)
 DT

DE Mouse activation-induced cytidine deaminase encoding cDNA SEQ ID NO:1.

XX Activation-induced cytidine deaminase; AID; cytidine deaminase;
 KW immune related disease; allergy; allergic disease; antiallergic;
 KW antianemic; antisthmatic; ophthalmological; anti-HIV; dermatological;
 KW gene therapy; B cell associated immune system disorder; food allergy;
 KW immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
 KW IGA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis;
 KW drug allergy; allergic rhinitis; Rosen disease; DiGeorge disease; AIDS;
 KW ataxia telangiectasia; common variable immunodeficiency disorder;
 KW major histocompatibility class II deficiency disease;
 KW auto immunodeficiency syndrome; IgG subclass selection disorder; ss.

XX Mus musculus.

XX Key Location/Qualifiers
 XX FH 93..689
 FT CDS /*tag= a
 FT /product= "activation-induced cytidine deaminase"

XX MO200058480-A1.
 XX 05-OCT-2000.

XX 28-MAR-2000; 2000WO-TP001918.

XX 29-MAR-1999; 99JP-00087192.
 XX 24-JUN-1999; 99JP-00178999.
 XX 27-DEC-1999; 99JP-00371382.

XX (NIBS) JAPAN TOBACCO INC.
 PA (HONJ) HONJO T.

XX Honjo T, Muramatsu M;
 XX WPI: 2000-611715/58.
 DR P-PSDB; AAB24197.

XX Nucleic acid encoding activation induced cytidine deaminase, useful as a
 PT target for drug development for immune-related diseases including
 PT allergies.

XX Claim 3; Page 126-130; 174pp; Japanese.

XX The present sequence encodes mouse activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic,
 CC antisthmatic, ophthalmological, anti-HIV and dermatological activities.

CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
 CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, rosacea disease, Disgeorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, MHC (major
 CC histocompatibility class) class II deficiency disease, AIDS (auto
 CC immunodeficiency syndrome), elevated IGE disorder, and IGE subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders
 CC
 SQ Sequence 2440 BP; 706 A; 546 C; 551 G; 636 T; 0 U; 1 Other;

Query Match 74.1%; Score 109.6; DB 3; Length 2440;
 Best Local Similarity 83.8%; Pred. No. 4e-29;
 Matches 124; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 CCTCTGATGAACCGAGAGAGTTCTTACCAATTCAAAATGTCGCGTGGGCTAAGG 60
 DB 101 CCTTCATGAGAGAAAGAGTTCTTACCAATTCAAAATGTCGCGTGGGCTAAGG 160
 QY 61 TCGGCGTGAAGCCTACCTGCTACCTAGTGAAGAGCGGTGACAGTCTACATCTTTTC 120
 DB 161 ACCGATGAGACCTACCTGCTACCTAGTGAAGAGAGATAGTCCACCTCTGCTC 220
 QY 121 ACTGAGCTTGTGTATCTTGGCAATAG 148
 DB 221 ACTGAGCTTGGCCACCTTGGCAACAG 248

RESULT 11

ABE80598
 ID ABL80598 standard; cDNA; 254 BP.

AC ABL80598;

DT 17-MAY-2002 (first entry)

DE Human ovarian cancer related cDNA clone SEQ ID NO:3576.

KM Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

OS Homo sapiens.

PN WO200192581-A2.

PD 06-DEC-2001.

PF 29-MAY-2001; 2001WO-US017756.

PR 26-MAY-2000; 2000US-0207484P.

PA (CORI-) CORIXA CORP.

PI Algate PA, Harlocker SL, Jones R;

DR WPI; 2002-122075/16.

PT Composition for therapy and diagnosis of ovarian cancer comprising

PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding

PT polypeptide, antibody specific to polypeptide or T cell expressing

PS Claim 1; SEQ ID NO 3576; 489pp; English.

XX The present invention describes a composition (I) comprising: carriers
 CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (SI)
 CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
 CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
 CC or antigen presenting cells that express (II). (I) has cytostatic

CC activity. An oligonucleotide (IV) that hybridises to (SI) can be used for
 CC detecting ovarian cancer in a patient's biological sample preferably
 CC serum or ovarian tissue. The method comprises contacting a biological
 CC sample from a patient with (IV), detecting the amount of polynucleotide
 CC hybridising to (IV) and comparing the amount to a predetermined cutoff
 CC value and thereby detecting ovarian cancer in the patient, where the
 CC amount of polynucleotide hybridising to (IV) is detected preferably by
 CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
 CC useful for stimulating and/or expanding T cells specific for an ovarian
 CC tumour protein comprising contacting T cells with (III) or (II). (III) is
 CC useful in design and preparation of ribozyme molecules for inhibiting
 CC expression of the tumour polypeptides and proteins in tumour cells; and
 CC to isolate a full length gene from a suitable library e.g., a tumour cDNA
 CC library using well known techniques
 CC
 SQ Sequence 254 BP; 57 A; 66 C; 65 G; 59 T; 0 U; 7 Other;

Query Match 27.8%; Score 41.2; DB 6; Length 254;
 Best Local Similarity 57.9%; Pred. No. 0.00012;
 Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 2 CTTGTGATGAACCGAGAGAGTTCTTACCAATTCAAAATGTCGCGTGGGCTAAGG 61
 DB 7 CACTGATGATCCACCCACATTCATTCAACTTAACATGACCTTGGGCTAGAGGA 66
 QY 62 CGGCGTGAAGCCTACCTGCTACCTAGTGAAGAGCGGTGACAGTCTACATCTTTTCA 121
 DB 67 CGGATGAGACTTACCTGCTGTATGAGTGAAGGCGCATGACATGACATGACCTGGTCTG 126
 QY 122 CTGAGC 127
 DB 127 CTGAAC 132

RESULT 12

AAA12409
 ID AAA12409 standard; cDNA; 1446 BP.

AC AAA12409;

DT 25-JUL-2000 (first entry)

DE cDNA encoding a human RNA-associated protein.

KM Human; RNA-associated protein; cell proliferation; cancer; inflammation;

KM immune response; reproductive disorder; actinic keratosis;

KM atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;

KM mixed connective tissue disease; myelofibrosis; primary thrombocythemia;

KM paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;

KM trauma; ss.

OS Homo sapiens.

PN WO200015799-A2.

PD 23-MAR-2000.

PT 17-SEP-1999; 99WO-US021688.

PR 17-SEP-1998; 98US-00156039.

PR 22-SEP-1998; 98US-00158720.

PR 04-NOV-1998; 98US-00186415.

PR 08-APR-1999; 99US-0128660P.

XX (INCY-) INCYTE PHARM INC.
 PI Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;
 PI Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;

XX WPI: 2000-271437/23.
 DR P-PSDB; AAY84437.
 XX
 PT New polypeptides and polynucleotides, useful for preventing and treating
 PT a disorder associated with increased or decreased expression of RNA
 PT associated proteins.
 XX
 PS Claim 9; Page 119; 131pp; English.
 XX
 CC The present sequence encodes a human RNA-associated protein. The
 CC expression of RNA-associated proteins is closely associated with
 CC reproductive tissues, nervous tissues, cell proliferation including
 CC cancer, inflammation and immune responses, and so they may be used for
 CC diagnosis, treatment or prevention of cell proliferative, diseases and
 CC immune/inflammatory disorders, and reproductive disorders. Diseases and
 CC disorders which may be treated include actinic keratosis,
 CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
 CC connective tissue disease, myelofibrosis, paroxysmal nocturnal
 CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia
 CC and cancers, and trauma
 XX
 SQ Sequence 1446 BP; 370 A; 374 C; 393 G; 308 T; 0 U; 1 Other;
 Query Match 27.8%; Score 41.2; DB 3; Length 1446;
 Best Local Similarity 57.9%; Pred. No. 0.00022;
 Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
 QY 2 CTCTTGATGAACCGAGAGAGTTCTTACCAATTCAAAATGTCGCTGGCTAAGGT 61
 DB 864 CACTCGATGATCCACCACTTACTTCACTTTACATTAACCTTGGTCAGAGGA 923
 QY 62 CGGCGTAGAGCTTACTGTGTAGCTAGTGAAGAGCGGTGACAGTGTACATCTTTCA 121
 DB 924 CGGATGAGACTTACTGTGTATAGGTGAGGCGCATGACATGACCTGGTCTTG 983
 QY 122 CTGAGC 127
 DB 984 CTGAGC 989
 RESULT 13
 AA220856
 ID AA220856 standard; cDNA; 1534 BP.
 XX
 AC AA220856;
 XX
 DT 09-DEC-1999 (first entry)
 XX
 DE Polynucleotide sequence of the 1p547_4 clone.
 XX
 KW secreted protein; cDNA library; clone; transmembrane protein;
 KM signal sequence cloning; hybridization cloning; gene therapy; receptor;
 KM ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 51..1205
 FT /*tag= a
 FT /product= "1p547_4"
 FT /*tag= b
 FT /*tag= 981..1205
 FT mat_peptide
 FT /*tag= c
 XX
 PN W09942470-A1.
 XX
 PD 26-AUG-1999.
 XX
 PF 18-FEB-1999; 99WO-US003458.
 XX
 PR 18-FEB-1998; 98US-0075038P.

PR 17-FEB-1999; 99US-00251600.
 XX
 PA (GENY) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavalie ER, Collins-Racie LA, Merberg D;
 PI Treacy M, Agostino MJ, Steininger RJ;
 XX
 DR WPI: 1999-518580/43.
 DR P-PSDB; AAY42383.
 XX
 PT New polynucleotides encoding human secreted proteins used for
 PT therapeutic, diagnostic and research purposes.
 XX
 PS Claim 16; Page 104; 125pp; English.
 XX
 CC This is the polynucleotide sequence of the clone 1p547_4, which was
 CC isolated from a human fetal brain cDNA library using methods which are
 CC selective for cDNAs encoding secreted proteins, or by identification as a
 CC secreted or transmembrane protein on the basis of computer analysis of
 CC the amino acid sequence of the encoded protein. The pNs and proteins of
 CC the invention are predicted to have biological activities which would
 CC make them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is given.
 CC Suggested activities include nutritional activity, cytokine and cell
 CC proliferation/differentiation activity, immune stimulating (e.g. as
 CC vaccines) or suppressing activity, hematopoiesis regulating activity,
 CC tissue growth activity, activin/inhibin activity, thrombolytic activity,
 CC chemotactic/chemokinetic activity, hemostatic and, thrombin/tumor
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumor
 CC invasion suppressor activity, and tumor inhibition activity. The pNs are
 CC also stated to be useful for gene therapy. Other activities include
 CC inhibiting the growth, infection or function of bacteria, fungi, viruses
 CC and other parasites; effecting bodily characteristics such as, e.g.
 CC weight, color, skin, etc., effecting biorythms or circadian cycles;
 CC enhancing fertility; treatment of depression; treatment of pain; hormonal
 CC or endocrine activity
 XX
 SQ Sequence 1534 BP; 495 A; 363 C; 345 G; 331 T; 0 U; 0 Other;
 Query Match 27.8%; Score 41.2; DB 2; Length 1534;
 Best Local Similarity 57.9%; Pred. No. 0.00022;
 Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
 QY 2 CTCTTGATGAACCGAGAGAGTTCTTACCAATTCAAAATGTCGCTGGCTAAGGT 61
 DB 633 CACTCGATGATCCACCACTTACTTCACTTTACATTAACCTTGGTCAGAGGA 692
 QY 62 CGGCGTAGAGCTTACTGTGTAGCTAGTGAAGAGCGGTGACAGTGTACATCTTTCA 121
 DB 693 CGGATGAGACTTACTGTGTATAGGTGAGGCGCATGACATGACATGACCTGGTCTTG 752
 QY 122 CTGAGC 127
 DB 753 CTGAGC 758
 RESULT 14
 AAS59293
 ID AAS59293 standard; cDNA; 1534 BP.
 XX
 AC AAS59293;
 XX
 DT 16-JAN-2002 (first entry)
 XX
 DE Human cDNA encoding a secreted protein 1p547_4.
 XX
 KW Human; secreted protein; sr; antiinflammatory; immunosuppressive;
 KW nootropic; neuroprotective; antiarthritic; antimicrobial; vlnetary;
 KW cytostatic; antidiabetic; virucide; antiinfectivity; anticonvulsant;
 KW vasotropic; antiparkinsonian; immunostimulant; dermatological;
 KW antineumatic; antitumor; antidiacer; osteopathic; tranquiliser;
 KW cerebroprotective; cytokine; cell proliferation; cell differentiation;
 KW immune deficiency; severe combined immunodeficiency; SCID; tumour;

KM autoimmune disorder; multiple sclerosis; rheumatoid arthritis;
 KM graft-versus-host disease; myeloid deficiency; wound healing; ulcer;
 KM periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease;
 KM Parkinson's disease; Huntington's disease; infection; cardiac disease;
 KM stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
 KM food supplement.
 OS Homo sapiens.
 XX MO200175068-A2.
 XX 11-OCT-2001.
 XX 22-MAR-2001; 2001WO-US009369.
 XX 30-MAR-2000; 2000US-00539330.
 XX 04-DEC-2000; 2000US-00729674.
 XX (GENY) GENETICS INST INC.
 PI Jacobs K, Mccoy JM, Lavallie B, Collins-Racie LA, Evans C;
 PI Treacy M, Agostino MJ, Steininger RJ, Spaulding V, Wong GG, Clark H;
 PI Fechtel K, Weidberg D;
 XX WPI; 2001-639363/73.
 XX P-PSDB; AAU39075.
 PT Secreted human proteins, useful as vaccine for treating various diseases
 PT such as autoimmune disorders (e.g. multiple sclerosis), and nervous
 PT system disorders (e.g. stroke).
 XX
 XX Disclosure; Page 580; 619pp; English.
 XX The invention relates to novel human secreted proteins, the nucleic acids
 CC encoding them. The protein may exhibit cytokine, cell proliferation or
 CC cell differentiation activity or may induce production of other cytokines
 CC in certain cell populations and may exhibit immune stimulating or immune
 CC suppressing activity, which is useful for the treatment of various immune
 CC deficiencies and disorders e.g. severe combined immunodeficiency (SCID),
 CC autoimmune disorders e.g. multiple sclerosis, systemic lupus
 CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation.
 CC The proteins are also useful in the treatment of diseases and disorders
 CC including tissue, skin and organ transplantation and in graft-versus-host
 CC diseases (GVHD), in the induction of tumour immunity, myeloid or lymphoid
 CC cell deficiencies, wound healing and tissue repair, in the treatment of
 CC burns, infections and ulcers; as well as in treatment of periodontal
 CC disease, osteoporosis or osteoarthritis, mediated by inflammatory
 CC processes, diseases of the peripheral nervous system, Alzheimer's,
 CC Parkinson's disease, Huntington's disease, amyotrophic lateral
 CC sclerosis, and Shy-Drager syndrome, infections, infarction of cardiac and
 CC central nervous system vessel e.g. stroke, sepsis, inflammatory bowel
 CC disease, ulcers, bone regeneration. The protein, having activin- or
 CC inhibin-related activities is useful as a contraceptive based on the
 CC ability of inhibiting to decrease fertility in female mammals and decrease
 CC spermatogenesis in male mammals. The proteins and nucleic acids are also
 CC useful as food supplements. The present sequence encodes a secreted
 CC protein of the invention.
 XX
 XX Sequence 1534 BP; 495 A; 363 C; 345 G; 331 T; 0 U; 0 Other;
 SQ
 Query Match 27.8%; Score 41.2; DB 4; Length 1534;
 Best Local Similarity 57.9%; Pred. No. 0.00022;
 Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
 QY 2 CTTTATGACCGAGGAGGTTCTTACCAATTCAAAATGCGCGGTAGGGT 61
 DB 633 CACTCGATGATCCACCACTTCACTTCACTTAACTTAACATTAACCTTGCGACAGGA 692
 QY 62 CGGGGTGAGACTTACCTGTGTGTCGTAGTGAAGAGCGGTGACAGTGTCAATCCTTTTGA 121
 DB 693 CGGATGAGACTTACCTGTGTGTATGAGTGAAGGAGCATGACATGACCTCGGCTCG 752
 QY 122 CTGAC 127

DB 753 CTGAC 758
 RESULT 15
 ID ABA90962 standard; cDNA, 1534 BP.
 XX ABA90962;
 AC ABA90962;
 XX 14-FEB-2002 (first entry)
 DT
 XX Human polynucleotide SEQ ID NO 173.
 DE Human; clone bd306-7; clone YB8-1; ATCC number 9859; gene therapy;
 KM immune disorder; bacterial infection; fungal infection; cancer; tumour;
 KM autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibin;
 KM osteoporosis; osteoarthritis; nervous system disorder; neuropathy;
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;
 KM haemophilia; cardiac infarction; stroke; sepsis; arthritis; vulnery;
 KM ischaemia-reperfusion injury; inflammatory bowel disease; chemotactic;
 KM Crohn's disease; cytosolic; anti-inflammatory; immunomodulator;
 KM neuroprotective; haemostatic; thrombolytic; anti-inflammatory; ss.
 XX Homo sapiens.
 OS
 XX US2001039335-A1.
 XX 08-NOV-2001.
 XX 04-DEC-2000; 2000US-00729674.
 XX 26-NOV-1997; 97US-0126425P.
 XX 04-DEC-1997; 97US-0067454P.
 XX 20-DEC-1997; 97US-0068379P.
 XX 02-JAN-1998; 98US-0070346P.
 XX 07-JAN-1998; 98US-0070643P.
 XX 08-JAN-1998; 98US-0070755P.
 XX 13-JAN-1998; 98US-0071304P.
 XX 22-JAN-1998; 98US-0072134P.
 XX 30-JAN-1998; 98US-0073095P.
 XX 18-FEB-1998; 98US-0075038P.
 XX 23-NOV-1998; 98US-00197886.
 XX 30-MAR-2000; 2000US-00539330.
 XX (JACO) JACOBS K.
 PA (MCCO) MCCOY J M.
 PA (LAVA) LAVALLIE B R.
 PA (COLL) COLLINS-RACIE L A.
 PA (EVAN) EVANS C.
 PA (MERB) MERBERG D.
 PA (TREAC) TREACY M.
 PA (AGOS) AGOSTINO M J.
 PA (STRI) STEININGER R J.
 PA (SPAU) SPAULDING V.
 PA (WONG) WONG G G.
 PA (CLAR) CLARK H.
 PA (FECH) FECHTEL K.
 XX
 XX Metcroy K, Mccoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
 PI Wong GG, Clark H, Fechtel K;
 XX WPI; 2002-040725/05.
 XX P-PSDB; ABB55784.
 XX New secreted proteins and encoding polynucleotides, useful in gene
 PT therapies, particularly for preventing or treating autoimmune disorders,
 PT cancer, graft-versus-host disease, wound, osteoporosis, stroke or
 PT inflammations.
 PS Disclosure; Page 310-311; 349pp; English.

CC The invention relates to isolated polynucleotides (ABA90876-ABA90968 and
 CC ABA90980) and encoded proteins (ABBS5698-ABBS5800), especially
 CC polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and
 CC proteins SEQ ID NO 2 (ABBS5698) and SEQ ID NO 20 (ABBS5707) contained in
 CC clones bd306-7 and yb8-1 respectively and the clones bd306-7 and yb8-1
 CC are deposited with the American Type Culture Collection (ATCC) with
 CC accession number 98599. The polynucleotides and encoded polypeptides have
 CC cytosolic, anti-inflammatory, immunomodulator, vulnery, neuroprotective, activin, inhibit, chemotactic, haemostatic, thrombolytic
 CC and anti-inflammatory activity and acting as cytokine modulators,
 CC haematopoiesis regulators, tissue growth modulators and/or cadherin
 CC suppressors. The polypeptides and polynucleotides are useful in gene
 CC therapies, particularly for preventing, treating or ameliorating any of
 CC the following diseases: immune deficiency and disorders; e.g. bacterial
 CC or fungal infections, autoimmune disorders, cancer, systemic lupus
 CC erythematosus or graft-versus-host disease; myeloid or lymphoid cell
 CC deficiencies; wound, burns, incisions and ulcers, osteoporosis or
 CC osteoarthritis; central and peripheral nervous system diseases and
 CC neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's
 CC disease; amyotrophic lateral sclerosis or Shy-Drager syndrome;
 CC haemophilia, cardiac infarction or stroke; inflammations, shock, sepsis
 CC or systemic inflammatory response syndrome, ischaemia-reperfusion injury,
 CC endotoxin lethality, arthritis, inflammatory bowel disease or Crohn's
 CC disease; or tumours or cancers, pemphigus vulgaris or pemphigus foliaceus
 XX

SQ Sequence 1534 BP; 495 A; 363 C; 345 G; 331 T; 0 U; 0 Other;

Query Match 27.8%; Score 41.2; DB 6; Length 1534;

Best Local Similarity 57.9%; Pred. No. 0.00022; Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| QY | 2 | CTCTTGATGAACCGAGAACTTCTTACCAATTCAAAATGTCGGCTAAGGT | 61 |
| DB | 633 | CACTCGATGATCCACCCACATTCACCTTCACTTAACAATGAACTTGGTCAGAGGA | 692 |
| QY | 62 | CGGCGTAGACCTACCTGTGTAGTAGTAAGAGGCGTACAGTGCTACATCCTTTCA | 121 |
| DB | 693 | CGGCATAGACTTACCTGTGTATAGAGTGAGCGCATGCACAATGACACTGGGTCTG | 752 |
| QY | 122 | CTGGAC | 127 |
| DB | 753 | CTGAAC | 758 |

Search completed: March 12, 2004, 18:08:07
 Job time : 61.2641 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 17:39:25 / Search time 11.019 Seconds

(without alignments)
7458.540 Million cell updates/sec

Title: US-09-966-880A-12

Perfect score: 148
Sequence: 1 cctctgacgaacccgagga.....ttggtatcttcgaataag 148

Scoring table: IDENTITY_NTC

Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Issued Patents NA:
1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
4: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
5: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
6: /cgn2_6/prodata/2/ina/5B_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------------------|-------------------|
| C 1 | 31.8 | 21.5 | 12222 | 4 US-09-328-925-42 | Sequence 42, Appl |
| C 2 | 27.8 | 18.8 | 3984 | 4 US-09-848-294-1 | Sequence 1, Appl |
| C 3 | 26.8 | 18.1 | 610 | 1 US-08-816-241-2 | Sequence 2, Appl |
| C 4 | 26.8 | 18.1 | 610 | 3 US-09-128-395-2 | Sequence 2, Appl |
| C 5 | 26.6 | 18.0 | 2208 | 4 US-09-543-681A-3435 | Sequence 3435, Ap |
| C 6 | 25.6 | 17.3 | 3586 | 3 US-08-847-296B-3 | Sequence 3, Appl |
| C 7 | 25.6 | 17.3 | 14561 | 4 US-09-392-714-1 | Sequence 1, Appl |
| C 8 | 25.4 | 17.2 | 8981 | 4 US-09-526-193A-20 | Sequence 20, Appl |
| C 9 | 25.4 | 17.2 | 62804 | 4 US-09-800-960-3 | Sequence 3, Appl |
| C 10 | 25.4 | 17.2 | 62804 | 4 US-10-096-960-3 | Sequence 3, Appl |
| C 11 | 25.2 | 17.0 | 4797 | 4 US-09-419-568F-25 | Sequence 25, Appl |
| C 12 | 25.2 | 17.0 | 4797 | 4 US-09-354-243B-25 | Sequence 25, Appl |
| C 13 | 25.2 | 16.9 | 696 | 4 US-08-252-991A-13591 | Sequence 13591, A |
| C 14 | 25.2 | 16.9 | 1077 | 4 US-08-252-991A-13599 | Sequence 13599, A |
| C 15 | 25.2 | 16.9 | 1467 | 4 US-08-252-991A-13960 | Sequence 13960, A |
| C 16 | 25.2 | 16.9 | 2049 | 4 US-09-252-991A-13850 | Sequence 13850, A |
| C 17 | 24.6 | 16.6 | 670 | 4 US-09-231-017B-569 | Sequence 569, App |
| C 18 | 24.6 | 16.6 | 864 | 3 US-08-924-570A-1 | Sequence 1, Appl |
| C 19 | 24.6 | 16.6 | 1082 | 4 US-09-247-155-181 | Sequence 181, App |
| C 20 | 24.6 | 16.6 | 1394 | 4 US-09-247-155-176 | Sequence 743, App |
| C 21 | 24.6 | 16.6 | 1526 | 4 US-09-221-017B-743 | Sequence 743, App |
| C 22 | 24.6 | 16.6 | 1613 | 4 US-08-965-048-3 | Sequence 3, Appl |
| C 23 | 24.6 | 16.6 | 1817 | 4 US-08-965-048-1 | Sequence 1, Appl |
| C 24 | 24.6 | 16.6 | 1857 | 4 US-09-620-312D-879 | Sequence 879, App |
| C 25 | 24.6 | 16.6 | 18596 | 3 US-09-318-448-11 | Sequence 11, Appl |
| C 26 | 24.6 | 16.6 | 18597 | 4 US-09-966-665-8 | Sequence 8, Appl |
| C 27 | 24.6 | 16.6 | 18597 | 4 US-09-963-333-8 | Sequence 8, Appl |

| | | | | | |
|------|------|------|---------|----------------------|-------------------|
| C 28 | 24.6 | 16.6 | 20966 | 4 US-09-964-880-3 | Sequence 3, Appl |
| C 29 | 24.6 | 16.6 | 20966 | 4 US-10-277-032-3 | Sequence 3, Appl |
| C 30 | 24.6 | 16.6 | 45716 | 4 US-08-965-048-5 | Sequence 5, Appl |
| C 31 | 24.6 | 16.6 | 45989 | 4 US-08-965-048-6 | Sequence 6, Appl |
| C 32 | 24.4 | 16.5 | 1021 | 4 US-09-221-017B-750 | Sequence 750, App |
| C 33 | 24.4 | 16.5 | 16075 | 3 US-09-096-942-1 | Sequence 1, Appl |
| C 34 | 24.4 | 16.5 | 16075 | 3 US-09-096-867-1 | Sequence 1, Appl |
| C 35 | 24.4 | 16.5 | 1830121 | 4 US-09-557-884-1 | Sequence 1, Appl |
| C 36 | 24.4 | 16.5 | 1830121 | 4 US-09-643-990A-1 | Sequence 1, Appl |
| C 37 | 24.4 | 16.5 | 4403765 | 3 US-09-103-840A-2 | Sequence 2, Appl |
| C 38 | 24.4 | 16.5 | 4411529 | 3 US-09-103-840A-1 | Sequence 1, Appl |
| C 39 | 23.8 | 16.1 | 420 | 3 US-09-060-756-417 | Sequence 417, App |
| C 40 | 23.8 | 16.1 | 420 | 4 US-09-670-314-417 | Sequence 417, App |
| C 41 | 23.8 | 16.1 | 588 | 1 US-08-742-023-4 | Sequence 4, Appl |
| C 42 | 23.8 | 16.1 | 588 | 1 US-08-742-023-5 | Sequence 5, Appl |
| C 43 | 23.8 | 16.1 | 588 | 3 US-08-968-505-4 | Sequence 4, Appl |
| C 44 | 23.8 | 16.1 | 588 | 3 US-08-968-505-5 | Sequence 5, Appl |
| C 45 | 23.8 | 16.1 | 2076 | 3 US-08-123-934A-3 | Sequence 3, Appl |

ALIGNMENTS

RESULT 1
US-09-328-925-42/c
Sequence 42, Application US/09328925
Patent No. 6610906
GENERAL INFORMATION:
APPLICANT: Kurachi, Kotoku
APPLICANT: Kurachi, Sumio
TITLE OF INVENTION: Nucleotide Sequences for Gene Regulation and Methods of
FILE REFERENCE: UM-03503
CURRENT FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 84
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 42
LENGTH: 12222
TYPE: DNA
ORGANISM: Homo sapiens
US-09-328-925-42

Query Match 21.5%, Score 31.8, DB 4; Length 12222;
Best Local Similarity 53.7%; Pred. No. 0.047;
Matches 66; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 18 GGAAGTTCTTACCAATTCAAATGTCGGCTGAGGCGGCGTGAACCTTACC 77
DB 202 GGCCTTTGTTTCCATCTGCACATGACCCCACTGAAGCCTGCTGTTCTCCAC 143
QY 78 TGTGTACGTGTGAGAGGCGGTGACAGTGTACATCTTTCACTGAGCTTGTATC 137
DB 142 TGTGAATGTAGTACGACGACGACGACATGTTCACCTTCCCTCTCCTCAAGCTC 83
QY 138 TTC 140
DB 82 CTC 80
RESULT 2
US-09-848-294-1/c
Sequence 1, Application US/09848294
Patent No. 6479640
GENERAL INFORMATION:
APPLICANT: Tonks, Nicholas K.
TITLE OF INVENTION: Isolation of A CDNA Encoding A No. 6479640el
TITLE OF INVENTION: Protein Tyrosine Phosphatase Which Localizes to Focal
FILE REFERENCE: CSHL90-04FZA
CURRENT APPLICATION NUMBER: US/09/848,294
CURRENT FILING DATE: 2001-05-03
PRIOR APPLICATION NUMBER: 09/235,251

PRIOR FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 08/759,536
PRIOR FILING DATE: 1996-12-04
PRIOR APPLICATION NUMBER: 08/107,420
PRIOR FILING DATE: 1993-08-16
PRIOR APPLICATION NUMBER: 07/663,579
PRIOR FILING DATE: 1991-03-01
PRIOR APPLICATION NUMBER: 07/494,036
PRIOR FILING DATE: 1990-03-14
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 3984
TYPE: DNA
ORGANISM: Homosapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (24)...(2765)
US-09-848-294-1

Query Match 18.8%; Score 27.8; DB 4; Length 3984;
Best Local Similarity 53.2%; Pred. No. 1.1;
Matches 59; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 27 TTTACCAATTCAAAATGTCCTGCTGAGGCTGAGACCTACCTGTGTAGC 86
DB 2500 TTTACCAATTCAAAATGTCGAGAGATCATCGGATACCGGTGAGCCATGCGACG 2441
QY 87 TAGTGAAGAGCGCGTGAAGTCTACATCTTTTCACTGACCTTGTATC 137
DB 2440 TAGTGAAGATGTGTCACTGTGTCTTCTCCCGGTGAGGTGTTGTAGCC 2390

RESULT 3

US-08-816-241-2
Sequence 2, Application US/08816241
Patent No. 5804185
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goll, Surya K.
TITLE OF INVENTION: NOVEL RNA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/816,241
FILING DATE: Filed Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0239 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSTUT09
CLONE: 1646833
US-08-816-241-2

Query Match 18.1%; Score 26.8; DB 1; Length 610;
Best Local Similarity 57.0%; Pred. No. 1;
Matches 49; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 8 ATGAACCGAGAGAGTTCTTACCAATTCGAGCTGAGGCTAGGGTCCGCT 67
DB 34 ATGTATCCAGGCACTTCTACTTCATTTAAACCTATGGAAGCAACGATCGAAC 93
QY 68 GAGACCTACCTGCTGCTAGTGA 93
DB 94 GAAACTGGCTGTCTCACCCTGGA 119

RESULT 4

US-09-128-395-2
Sequence 2, Application US/09128395
Patent No. 6087108
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goll, Surya K.
TITLE OF INVENTION: NOVEL RNA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/128,395
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/816,241
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0239 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSTUT09
CLONE: 1646833
US-09-128-395-2

Query Match 18.1%; Score 26.8; DB 3; Length 610;
Best Local Similarity 57.0%; Pred. No. 1;
Matches 49; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 8 ATGAACCGAGAGAGTTCTTACCAATTCGAGCTGAGGCTAGGGTCCGCT 67
DB 34 ATGTATCCAGGCACTTCTACTTCATTTAAACCTATGGAAGCAACGATCGAAC 93

QY 68 GAGACCTACCTGTGCTACGTAGTAA 93
DB 94 GAACTTGCTGTGCTTACCGTGA 119

RESULT 5

US-09-543-681A-3435
; Sequence 3435, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709 1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3435
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-3435

Query Match 18.0%; Score 26.6; DB 4; Length 2208;
Best Local Similarity 56.2%; Pred. No. 2.4;
Matches 50; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 6 TGATGACCGAGAGATTCTTACCAATTCGCGCTGAGCGTAAAGCGTGGC 65
DB 498 TCAAAATATAGGAACTTCATTAGGATTCAAAATATAGCTGTACACTTACAG 557
QY 66 GTGAGACCTACCTGTGCTACGTAGTAA 94
DB 558 GAGACTGGAGTGTACTTAATAGTGAAG 586

RESULT 6

US-08-847-296B-3
; Sequence 3, Application US/08847296B
; Patent No. 6271347
; GENERAL INFORMATION:
; APPLICANT: DAUGHERTY, BRUCE L.
; APPLICANT: DEMARTINO, JULIE A.
; APPLICANT: SICILIANO, SALVATORE J.
; APPLICANT: SPRINGER, MARTIN J.
; TITLE OF INVENTION: EOSINOPHIL EOTAXIN RECEPTOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,296B
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/016,158
; FILING DATE: 26-APR-1996
; APPLICATION NUMBER: 60/017,113
; FILING DATE: 26-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Thies, J. Eric
; REGISTRATION NUMBER: 35,382

REFERENCE/DOCKET NUMBER: 19634Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-3904
; TELEFAX: 908-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3586 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-847-296B-3

Query Match 17.3%; Score 25.6; DB 3; Length 3586;
Best Local Similarity 50.8%; Pred. No. 7.4;
Matches 61; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 21 AGTTCTTACCAATTCAAAATGTCGCTGAGGCTGAGAGCTTACCTGT 80
DB 1893 AGTCATTCTTCTCTCTCAAAAGCTGGAATGTGCTATGATGAGAGATTA 1952
QY 81 GCTACGTGAGAGAGCGGTGACAGTGTCTACATCTTTTCACTGACTTGTATCTTC 140
DB 1953 ACAGACCATGATAAAGAGATCAACAAGTCCACCAGGAGCCCTATTTTCTAATTT 2012

RESULT 7

US-09-392-714-1/c
; Sequence 1, Application US/09392714A
; Patent No. 6666147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; PRIOR FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 14561
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-392-714-1

Query Match 17.3%; Score 25.6; DB 4; Length 14561;
Best Local Similarity 55.7%; Pred. No. 15;
Matches 49; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 39 AAAATGTCGCTGAGGCTAAGGCTGAGACTGCTGTAGCTAGTGAAGAGGC 98
DB 9032 AAGACGAGAGAGAGAGAGAGAGCTGGAATGTGAGCTGAGACTGGAAGCGCTAAGGC 8973
QY 99 GTGACAGTGTACATCTTTTCACTGGA 126
DB 8972 GTGACACTGTACTTGTGTGCAAGGA 8945

RESULT 8

US-09-526-193A-20/c
; Sequence 20, Application US/09526193A
; Patent No. 6617122
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael R.
; APPLICANT: Brooks-Wilson, Angela R.
; APPLICANT: Pimstone, Simon N.

;; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
;; TITLE OF INVENTION: CHOLESTEROL LEVELS
;; FILE REFERENCE: 50110/002005
;; CURRENT APPLICATION NUMBER: US/09/526,193A
;; CURRENT FILING DATE: 2000-03-15
;; PRIOR APPLICATION NUMBER: 60/124,702
;; PRIOR FILING DATE: 1999-03-15
;; PRIOR APPLICATION NUMBER: 60/138,048
;; PRIOR FILING DATE: 1999-06-08
;; PRIOR APPLICATION NUMBER: 60/139,600
;; PRIOR FILING DATE: 1999-06-17
;; PRIOR APPLICATION NUMBER: 60/151,977
;; PRIOR FILING DATE: 1999-09-01
;; NUMBER OF SEQ ID NOS: 287
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 20
;; LENGTH: 8981
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-526-193A-20

Query Match 17.2%; Score 25.4; DB 4; Length 8981;
Best Local Similarity 61.2%; Pred. No. 14;
Matches 41; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 44 GTCCGCTGGGCTAGAGGCTGGCGGAGACCTACTGCTGCTACGTAGAGAGGCGGTGAC 103
DB 7558 GTCTGGTGGGCAATGGACATGTAGACATCTAACGCTGTACAGAGGAGATATACAC 7599
QY 104 AGTGCTA 110
DB 7598 AGGCCAA 7592

RESULT 9
US-09-800-960-3
;; Sequence 3, Application US/09800960
;; Patent No. 6387677
;; GENERAL INFORMATION:
;; APPLICANT: YE, Jane et al.
;; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;; TITLE OF INVENTION: THEREOF
;; FILE REFERENCE: CI001158
;; CURRENT APPLICATION NUMBER: US/09/800,960
;; CURRENT FILING DATE: 2001-03-08
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3
;; LENGTH: 62804
;; TYPE: DNA
;; ORGANISM: Human
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(62804)
;; OTHER INFORMATION: n = A,T,C or G
US-09-800-960-3

Query Match 17.2%; Score 25.4; DB 4; Length 62804;
Best Local Similarity 58.7%; Pred. No. 36;
Matches 44; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 19 GAAGTTTCTTACCAATCAAAAATGTCGCTGAGGCTAGAGGCTGAGACCTACT 78
DB 24634 GAATTTCTTAGGATTCGCCAATGTCCCTGCGGTGGGGGGGGGGGCAATTC 24693
QY 79 GTGCTAGCTAGTGA 93
DB 24694 ATCCCACTTGA 24708

RESULT 10
US-10-096-960-3

;; Sequence 3, Application US/10096960
;; Patent No. 6664085
;; GENERAL INFORMATION:
;; APPLICANT: YE, Jane et al.
;; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;; TITLE OF INVENTION: THEREOF
;; FILE REFERENCE: CI001158D1V
;; CURRENT APPLICATION NUMBER: US/10/096,960
;; CURRENT FILING DATE: 2002-03-14
;; PRIOR APPLICATION NUMBER: 09/800,960
;; PRIOR FILING DATE: 2001-03-08
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3
;; LENGTH: 62804
;; TYPE: DNA
;; ORGANISM: Homo sapien
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(62804)
;; OTHER INFORMATION: n = A,T,C or G
US-10-096-960-3

Query Match 17.2%; Score 25.4; DB 4; Length 62804;
Best Local Similarity 58.7%; Pred. No. 36;
Matches 44; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 19 GAAGTTTCTTACCAATCAAAAATGTCGCTGAGGCTAGAGGCTGAGACCTACT 78
DB 24634 GAATTTCTTAGGATTCGCCAATGTCCCTGCGGTGGGGGGGGGGGCAATTC 24693
QY 79 GTGCTAGCTAGTGA 93
DB 24694 ATCCCACTTGA 24708

RESULT 11
US-09-419-568F-25
;; Sequence 25, Application US/09419568F
;; Patent No. 6331613
;; GENERAL INFORMATION:
;; APPLICANT: Dumoutier, Laure
;; APPLICANT: Renauld, Jean-Christophe
;; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fa
;; TITLE OF INVENTION: (TfPs) The Proteins Encoded, and Uses Thereof
;; FILE REFERENCE: LUD 5543.2
;; CURRENT APPLICATION NUMBER: US/09/419,568F
;; CURRENT FILING DATE: 1999-10-18
;; PRIOR APPLICATION NUMBER: US09/354,243
;; PRIOR FILING DATE: 1999-07-16
;; PRIOR APPLICATION NUMBER: US09/178,973
;; PRIOR FILING DATE: 1998-10-26
;; NUMBER OF SEQ ID NOS: 29
;; SEQ ID NO 25
;; LENGTH: 4797
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
US-09-419-568F-25

Query Match 17.0%; Score 25.2; DB 4; Length 4797;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 36; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 87 TAGTGAAGAGGCGGTACAGCTGCTACATCTTTTCACTGGAATTTGGTATCTTC 140
DB 3750 TAGATTAGAGGAATTCAGTAGGGAATGCTTTTCACTTGAATTTGGGTTCTTC 3803

RESULT 12
US-09-354-243B-25

; Sequence 25, Application US/09354243B
; Patent No. 6359117
; GENERAL INFORMATION:
; APPLICANT: Dumoutier, Laure
; APPLICANT: Louned, Jamila
; APPLICANT: Renaud, Jean-Christophe
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fa
; TITLE OF INVENTION: (Tifs)
; FILE REFERENCE: LUD 5543.1
; CURRENT APPLICATION NUMBER: US/09/354,243B
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: US09/178,973
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 29
; SEQ ID NO 25
; LENGTH: 4797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-354-243B-25

Query Match 17.0%; Score 25.2; DB 4; Length 4797;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 36; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 87 TAGTGAAGGCGGTGACATGCTTCTTCACTGAGACTTGGTATCTTC 140
DB 3750 TAGATTAGAGAAATTTCAGTAGGAATGCTTTTCACTGAATTGGGTTCTC 3803

RESULT 13
US-09-252-991A-13391
; Sequence 13391, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13391
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13391

Query Match 16.9%; Score 25; DB 4; Length 696;
Best Local Similarity 49.6%; Pred. No. 5.8;
Matches 64; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 1 CCTCTGATGACCGGAGAGATTCTTTACCAATTCAAAATGTCGGCTGGGCTAAGG 60
DB 393 CCTGTTGCTGATGCTGGCTGGTGTGCACTGCAATCGGCCGACGCCGACGCTAAGT 452
QY 61 TCGGCGTGAAGACCTTACTGCTAGCTAGTGAAGAGCGGTGACAGTCTTCTTTTC 120
DB 453 GTTGCTGAAGCGCTTGTGAGCGTTGCGCGGAGCGCGGAGCGCGGCGCTTGT 512
QY 121 ACTGGAATT 129
DB 513 CTTGGCGTT 521

RESULT 14
US-09-252-991A-13599
; Sequence 13599, Application US/09252991A

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13599
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13599

Query Match 16.9%; Score 25; DB 4; Length 1077;
Best Local Similarity 49.6%; Pred. No. 7.1;
Matches 64; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 1 CCTCTGATGACCGGAGAGATTCTTTACCAATTCAAAATGTCGGCTGGGCTAAGG 60
DB 31 CCTGTTGCTGATGCTGGCTGGTGTGCACTGCAATCGGCCGACGCCGACGCTAAGT 90
QY 61 TCGGCGTGAAGACCTTACTGCTAGCTAGTGAAGAGCGGTGACAGTCTTCTTTTC 120
DB 91 GTTGCTGAAGCGCTTGTGAGCGTTGCGCGGAGCGCGGAGCGCGGCGCTTGT 150

QY 121 ACTGGAATT 129
DB 151 CTTGGCGTT 159

RESULT 15
US-09-252-991A-13960/C
; Sequence 13960, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13960
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13960

Query Match 16.9%; Score 25; DB 4; Length 1467;
Best Local Similarity 49.6%; Pred. No. 8.3;
Matches 64; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 1 CCTCTGATGACCGGAGAGATTCTTTACCAATTCAAAATGTCGGCTGGGCTAAGG 60
DB 1244 CCTGTTGCTGATGCTGGCTGGTGTGCACTGCAATCGGCCGACGCCGACGCTAAGT 1185
QY 61 TCGGCGTGAAGACCTTACTGCTAGCTAGTGAAGAGCGGTGACAGTCTTCTTTTC 120
DB 1184 GTTGCTGAAGCGCTTGTGAGCGTTGCGCGGAGCGCGGAGCGCGGCGCTTGT 1125
QY 121 ACTGGAATT 129
DB 1124 CTTGGCGTT 1116

Mon Mar 15 09:27:54 2004

us-09-966-880a-12.rn1

Page 6

Search completed: March 13, 2004, 00:51:14
Job time : 19.0119 secs

QY 121 ACTGACTTGTATCTTCCCAATAG 148
DB 121 ACTGACTTGTATCTTCCCAATAG 148

RESULT 2

US-09-966-880A-7
Sequence 7, Application US/09966880A
Patent No. US20020164743A1

GENERAL INFORMATION:

APPLICANT: Honjo, Tasuku

APPLICANT: Muramatsu, Masamichi

TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE

FILE REFERENCE: 06501-088001

CURRENT APPLICATION NUMBER: US/09/966,880A

PRIOR FILING DATE: 2001-09-28

PRIOR FILING DATE: 2000-03-28

PRIOR APPLICATION NUMBER: PCT/JP00/01918

PRIOR FILING DATE: 1999-12-27

PRIOR APPLICATION NUMBER: JP 11-178999

PRIOR FILING DATE: 1999-06-24

PRIOR APPLICATION NUMBER: JP 11-87192

PRIOR FILING DATE: 1999-03-29

NUMBER OF SEQ ID NOS: 36

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 2818

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (80)...(673)

FEATURE:

NAME/KEY: 5'UTR

LOCATION: (1)...(79)

FEATURE:

NAME/KEY: 3'UTR

LOCATION: (677)...(2818)

US-09-966-880A-7

Query Match 100.0%; Score 148; DB 9; Length 2818;
Best Local Similarity 100.0%; Pred. No. 1.8e-45;

Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTGATGACCGGAGAGTTCTTTACCAATCAAAATGTCGGCTGGCTAAAGG 60
DB 88 CCTTTGATGACCGGAGAGTTCTTTACCAATCAAAATGTCGGCTGGCTAAAGG 147

QY 61 TCGGCGTAGACCTACCTGCTGCTACGTAGTGAAGAGCGGTGACATCTTTTC 120
DB 148 TCGGCGTAGACCTACCTGCTGCTACGTAGTGAAGAGCGGTGACATCTTTTC 207

QY 121 ACTGACTTGTATCTTCCCAATAG 148
DB 208 ACTGACTTGTATCTTCCCAATAG 235

RESULT 3

US-09-966-880A-10

Sequence 10, Application US/09966880A

Patent No. US20020164743A1

GENERAL INFORMATION:

APPLICANT: Honjo, Tasuku

APPLICANT: Muramatsu, Masamichi

TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE

FILE REFERENCE: 06501-088001

CURRENT APPLICATION NUMBER: US/09/966,880A

PRIOR FILING DATE: 2001-09-28

PRIOR APPLICATION NUMBER: PCT/JP00/01918

PRIOR FILING DATE: 2000-03-28

PRIOR APPLICATION NUMBER: JP 11-371382

PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 6564
TYPE: DNA
ORGANISM: Homo sapiens
US-09-966-880A-10

Query Match 100.0%; Score 148; DB 9; Length 6564;

Best Local Similarity 100.0%; Pred. No. 2.6e-45;

Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTGATGACCGGAGAGTTCTTTACCAATCAAAATGTCGGCTGGCTAAAGG 60
DB 1065 CCTTTGATGACCGGAGAGTTCTTTACCAATCAAAATGTCGGCTGGCTAAAGG 1124

QY 61 TCGGCGTAGACCTACCTGCTGCTACGTAGTGAAGAGCGGTGACATCTTTTC 120
DB 1125 TCGGCGTAGACCTACCTGCTGCTACGTAGTGAAGAGCGGTGACATCTTTTC 1184

QY 121 ACTGACTTGTATCTTCCCAATAG 148
DB 1185 ACTGACTTGTATCTTCCCAATAG 1212

RESULT 4

US-09-966-880A-35

Sequence 35, Application US/09966880A

Patent No. US20020164743A1

GENERAL INFORMATION:

APPLICANT: Honjo, Tasuku

APPLICANT: Muramatsu, Masamichi

TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE

FILE REFERENCE: 06501-088001

CURRENT APPLICATION NUMBER: US/09/966,880A

PRIOR FILING DATE: 2001-09-28

PRIOR APPLICATION NUMBER: PCT/JP00/01918

PRIOR FILING DATE: 2000-03-28

PRIOR APPLICATION NUMBER: JP 11-371382

PRIOR FILING DATE: 1999-12-27

PRIOR APPLICATION NUMBER: JP 11-178999

PRIOR FILING DATE: 1999-06-24

PRIOR APPLICATION NUMBER: JP 11-87192

PRIOR FILING DATE: 1999-03-29

NUMBER OF SEQ ID NOS: 36

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 35

LENGTH: 11204

TYPE: DNA

ORGANISM: Homo sapiens

US-09-966-880A-35

Query Match 100.0%; Score 148; DB 9; Length 11204;

Best Local Similarity 100.0%; Pred. No. 3.3e-45;

Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTGATGACCGGAGAGTTCTTTACCAATCAAAATGTCGGCTGGCTAAAGG 60
DB 6280 CCTTTGATGACCGGAGAGTTCTTTACCAATCAAAATGTCGGCTGGCTAAAGG 6339

QY 61 TCGGCGTAGACCTACCTGCTGCTACGTAGTGAAGAGCGGTGACATCTTTTC 120
DB 6340 TCGGCGTAGACCTACCTGCTGCTACGTAGTGAAGAGCGGTGACATCTTTTC 6399

QY 121 ACTGACTTGTATCTTCCCAATAG 148
DB 6400 ACTGACTTGTATCTTCCCAATAG 6427

```

      SEQ ID NO 3576
      LENGTH: 254
      TYPE: DNA
      ORGANISM: Homo sapien
      FEATURE:
      NAME/KEY: misc_feature
      LOCATION: (1)..(254)
      OTHER INFORMATION: n = A,T,C or G
      US-09-667-701-3576

Query Match      27.8%; Score 41.2; DB 9; Length 254;
Best Local Similarity 57.9%; Pred. No. 3e-05;
Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY      2      CTCCTGATGAGACCGGAGGAACTTCTTTACCAATTCAAAATGTCGGCTGGGCTAAGGT 61
DB      7      CACTCGATGATGCCACCCCACTTCATTCTTAACATTAAGAACTTGGGTGAGAGA 66
      62      CGGGGTAGAGACCTGCTGTGTCTAGCTAGTGAAGAGCGCTGACAGTGTACATCCTTTTCA 121
      67      CGGATGAGACTTACTCTGTGTATTAGAGTGGAGCGCATGCACAAAGACCTGGGTCTCG 126

QY      122      CTGGAC 127
      127      CTGAC 132
DB      127      CTGAC 132

RESULT 7
US-09-796-692-4007
; Sequence 4007, Application US/09796692
; Publication No. US20020198362A1
GENERAL INFORMATION:
; APPLICANT: Galger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mammon, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796.692
PRIOR FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4007
LENGTH: 466
TYPE: DNA
ORGANISM: Homo sapiens
US-09-796-692-4007

Query Match      27.8%; Score 41.2; DB 9; Length 466;

```


Best Local Similarity 57.9%; Pred. No. 4e-05;
Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 2 CTCCTGATGAACCGAGAGAGTTCTTTTACCAATTCAAAAATGTCGCTGGCTAAGGT 61
DB 270 CACTCGATGATGATCCACCATTCCTTCACTTTAAACAATGAACCTTGGGTGAGAGA 329
QY 62 CGCGCTGAGACCTTACTGTGCTACGTAGTAAGAGCGGTGACAGTCTTCACTTTTCA 121
DB 330 CGGCATGAGACTTACCTGTGTATGAGTGAGCGCATGACACATGACACTGGGTCTG 389
QY 122 CTGGAC 127
DB 390 CTGAAC 395

RESULT 8

US-10-040-862-4007
; Sequence 4007, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane A.
; APPLICANT: Mannion, Jane
; APPLICANT: Reiter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4007
; LENGTH: 466
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-4007

Query Match 27.8%; Score 41.2; DB 14; Length 466;
Best Local Similarity 57.9%; Pred. No. 4e-05;
Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 2 CTCCTGATGAACCGAGAGAGTTCTTTTACCAATTCAAAAATGTCGCTGGCTAAGGT 61
DB 270 CACTCGATGATGATCCACCATTCCTTCACTTTAAACAATGAACCTTGGGTGAGAGA 329

QY 62 CGCGCTGAGACCTTACTGTGCTACGTAGTAAGAGCGGTGACAGTCTTCACTTTTCA 121
DB 330 CGGCATGAGACTTACCTGTGTATGAGTGAGCGCATGACACATGACACTGGGTCTG 389

QY 122 CTGGAC 127
DB 390 CTGAAC 395

RESULT 9

US-10-057-475B-4007
; Sequence 4007, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therap
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4007
; LENGTH: 466
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-4007

Query Match 27.8%; Score 41.2; DB 15; Length 466;
Best Local Similarity 57.9%; Pred. No. 4e-05;
Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 2 CTCCTGATGAACCGAGAGAGTTCTTTTACCAATTCAAAAATGTCGCTGGCTAAGGT 61
DB 270 CACTCGATGATGATCCACCATTCCTTCACTTTAAACAATGAACCTTGGGTGAGAGA 329
QY 62 CGCGCTGAGACCTTACTGTGCTACGTAGTAAGAGCGGTGACAGTCTTCACTTTTCA 121
DB 330 CGGCATGAGACTTACCTGTGTATGAGTGAGCGCATGACACATGACACTGGGTCTG 389

QY 122 CTGGAC 127
DB 390 CTGAAC 395

```
RESULT 10
US-10-154-884B-4007
; Sequence 4007, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4007
; LENGTH: 466
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-4007

Query Match          27.8%; Score 41.2; DB 15; Length 466;
Best Local Similarity 57.9%; Pred. No. 4e-05;
Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 2 CTCTTGATGAACCGAGAGAGTTCTTTACCAATTCAAAATGTCGCGTGGGCTAAGGCT 61
DB 270 CACTCGATGATGATCCACCCACATTCCTTCAACTTTAAATGATGACCTTGGGTCAAGAGA 329

QY 62 CGCGGTGAGACCTTACTGTGCTAGCTAGTGAAGAGCGGTGACAGTGTCTAATCCTTTTCA 121
DB 330 CGGATGAGACTTACTGTGTATGAGGTGAGCGCATGACATGACACTGGGTCTCTG 389

QY 122 CTGGAC 127
DB 390 CTGAAC 395

RESULT 11
US-09-796-692-4443
; Sequence 4443, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 2077.001200
```

```
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4443
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-4443

Query Match          27.8%; Score 41.2; DB 9; Length 505;
Best Local Similarity 57.9%; Pred. No. 4.1e-05;
Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 2 CTCTTGATGAACCGAGAGAGTTCTTTACCAATTCAAAATGTCGCGTGGGCTAAGGCT 61
DB 270 CACTCGATGATGATCCACCCACATTCCTTCAACTTTAAATGATGACCTTGGGTCAAGAGA 329

QY 62 CGCGGTGAGACCTTACTGTGCTAGCTAGTGAAGAGCGGTGACAGTGTCTAATCCTTTTCA 121
DB 330 CGGATGAGACTTACTGTGTATGAGGTGAGCGCATGACATGACACTGGGTCTCTG 389

QY 122 CTGGAC 127
DB 390 CTGAAC 395

RESULT 12
US-10-040-862-4443
; Sequence 4443, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
```

```
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: US 60/223,416
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR APPLICATION NUMBER: US 60/223,378
/ PRIOR FILING DATE: 2000-08-07
/ PRIOR APPLICATION NUMBER: US 09/796,692
/ PRIOR FILING DATE: 2001-03-01
/ NUMBER OF SEQ ID NOS: 10467
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 4443
/ LENGTH: 505
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-040-862-4443

Query Match      27.8% Score 41.2; DB 14; Length 505;
Best Local Similarity 57.9%; Pred. No. 4.1e-05;
Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 2 CTCTGATGAACCGAGAGAGTTCTTTACCAATTCAAAATGTCGCTGGCTAAGGT 61
   |||||
Db 270 CACTGATGATGCCACCCACATTCATTCACTTAAATGAACCTGGGTCAGAGA 329
   |||||

QY 62 CGGCGTGAAGACTTACTGTGTCTACGTAGTGAAGAGGGGTGCTATACCTTTTCA 121
   |||||
Db 330 CGGATGAGACTTACTGTGTGTATAGAGGTGAGGCGCATGCATGACACTGGGTCTTG 389
   |||||

QY 122 CTGAGC 127
   |||||
Db 390 CTGAAC 395

RESULT 13
US-10-057-475B-4443
/ Sequence 4443, Application US/10057475B
/ Publication No. US20040002068A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Clapper, Jonathan David
/ APPLICANT: Wang, Aijun
/ APPLICANT: Ordonez, Nadia
/ APPLICANT: Carter, Lauren
/ APPLICANT: McNeill, Patricia Dianne
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ TITLE OF INVENTION: Hematological Malignancies
/ FILE REFERENCE: 014058-01440205
/ CURRENT APPLICATION NUMBER: US/10/057,475B
/ CURRENT FILING DATE: 2002-01-22
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
```

```
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 10979
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 4443
/ LENGTH: 505
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-057-475B-4443

Query Match      27.8% Score 41.2; DB 15; Length 505;
Best Local Similarity 57.9%; Pred. No. 4.1e-05;
Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 2 CTCTGATGAACCGAGAGAGTTCTTTACCAATTCAAAATGTCGCTGGCTAAGGT 61
   |||||
Db 270 CACTGATGATGCCACCCACATTCATTCACTTAAATGAACCTGGGTCAGAGA 329
   |||||

QY 62 CGGCGTGAAGACTTACTGTGTCTACGTAGTGAAGAGGGGTGCTATACCTTTTCA 121
   |||||
Db 330 CGGATGAGACTTACTGTGTGTATAGAGGTGAGGCGCATGCATGACACTGGGTCTTG 389
   |||||

QY 122 CTGAGC 127
   |||||
Db 390 CTGAAC 395

RESULT 14
US-10-154-884B-4443
/ Sequence 4443, Application US/10154884B
/ Publication No. US2004000561A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therap
/ TITLE OF INVENTION: Hematological Malignancies
/ FILE REFERENCE: 014058-01352105
/ CURRENT APPLICATION NUMBER: US/10/154,884B
/ CURRENT FILING DATE: 2002-05-23
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 11290
```

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4443
LENGTH: 505
TYPE: DNA
ORGANISM: Homo sapiens
US-10-154-884B-4443

Query Match 27.8%; Score 41.2; DB 15; Length 505;
Best Local Similarity 57.9%; Pred. No. 4.1e-05;
Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 2 CTCTTGATGAACCGAGAGAGTTCTTTACCAATTCAAAAATGTCGCTGCGCTAAGGGT 61
Db 270 CACTCGATGATCCACCCACATTCATTCACTTAACAATGAACCTTGGGTCAAGGA 329
QY 62 CGGGGTGAGACTCTGTGCTAGTAGTGAAGGCGGTGACAGTGTACATCTTTCA 121
Db 330 CGGCATGAGACTTCTGTGTATGAGGTGAGCGCATGCAATGACACCTGGGTCTG 389
QY 122 CTGGAC 127
Db 390 CTGAAC 395

RESULT 15

US-10-460-923-1
Sequence 1, Application US/10460923
Publication No. US2004009951A1
GENERAL INFORMATION:
APPLICANT: MALIM, Michael H.
APPLICANT: SHEEHY, Ann M.
APPLICANT: HARRIS, Reuben S.
APPLICANT: BISHOP, Kate N.
APPLICANT: NEUBERGER, Michael S.
APPLICANT: GADDIS, Nathan C.
APPLICANT: SIMON, James H.M.
TITLE OF INVENTION: DNA Deamination Mediates Innate Immunity to Retroviral Infection
FILE REFERENCE: 22253-74380
CURRENT APPLICATION NUMBER: US/10/460,923
CURRENT FILING DATE: 2003-06-13
PRIOR APPLICATION NUMBER: US 60/388,513
PRIOR FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: US 60/472,952
PRIOR FILING DATE: 2003-05-23
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1155
TYPE: DNA
ORGANISM: Homo sapiens
US-10-460-923-1

Query Match 27.8%; Score 41.2; DB 15; Length 1155;
Best Local Similarity 57.9%; Pred. No. 5.8e-05;
Matches 73; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 2 CTCTTGATGAACCGAGAGAGTTCTTTACCAATTCAAAAATGTCGCTGCGCTAAGGGT 61
Db 583 CACTCGATGATCCACCCACATTCATTCACTTAACAATGAACCTTGGGTCAAGGA 642
QY 62 CGGGGTGAGACTCTGTGCTAGTAGTGAAGGCGGTGACAGTGTACATCTTTCA 121
Db 643 CGGCATGAGACTTCTGTGTATGAGGTGAGCGCATGCAATGACACCTGGGTCTG 702
QY 122 CTGGAC 127
Db 703 CTGAAC 708

Search completed: March 13, 2004, 05:30:56
Job time : 53.9361 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 13:19:34 ; Search time 471.269 Seconds
(without alignments)
9378.080 Million cell updates/sec

Title: US-09-966-880A-12

Perfect score: 148

Sequence: 1 ccccttgatcgaccgagga.....ttggttctcttcgcaataag 148

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estin:*
3: em_estin:*
4: em_estin:*
5: em_estin:*
6: em_estin:*
7: em_estin:*
8: em_estin:*
9: gb_est1:*
10: gb_est1:*
11: gb_est1:*
12: gb_est1:*
13: gb_est1:*
14: gb_est1:*
15: em_estfun:*
16: em_estfun:*
17: em_gse_hum:*
18: em_gse_hum:*
19: em_gse_hum:*
20: em_gse_hum:*
21: em_gse_hum:*
22: em_gse_hum:*
23: em_gse_hum:*
24: em_gse_hum:*
25: em_gse_hum:*
26: em_gse_hum:*
27: em_gse_hum:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|-------|-------------|-----------|----|--------------------|
| 1 | 148 | 100.0 | 541 | 10 | BF238155 601811880 |
| 2 | 148 | 100.0 | 743 | 12 | BG686133 602638412 |
| 3 | 148 | 100.0 | 820 | 12 | BG757089 602715124 |
| 4 | 148 | 100.0 | 872 | 12 | BG758510 602712721 |

| Result | Score | Query Match | Length DB | ID | Description |
|--------|-------|-------------|-----------|----|--------------------|
| 5 | 148 | 100.0 | 889 | 12 | BG686876 602650861 |
| 6 | 148 | 100.0 | 953 | 13 | B0065440 AGENCOURT |
| 7 | 148 | 100.0 | 1052 | 13 | B0055935 AGENCOURT |
| 8 | 148 | 100.0 | 1201 | 9 | AL559877 |
| 9 | 148 | 100.0 | 1201 | 3 | EX402063 |
| 10 | 146.4 | 98.9 | 535 | 14 | CD707143 |
| 11 | 137 | 92.6 | 693 | 12 | BG757392 |
| 12 | 137 | 92.6 | 942 | 10 | BF975166 |
| 13 | 136 | 91.9 | 853 | 13 | BX464579 |
| 14 | 117.6 | 79.5 | 701 | 28 | BZ247020 |
| 15 | 93.6 | 63.2 | 623 | 9 | AJ450295 |
| 16 | 93.6 | 63.2 | 650 | 9 | AJ449745 |
| 17 | 93.6 | 63.2 | 688 | 9 | AJ450317 |
| 18 | 93.6 | 63.2 | 696 | 9 | AJ453647 |
| 19 | 93.6 | 63.2 | 729 | 9 | AJ450296 |
| 20 | 93.6 | 63.2 | 757 | 9 | AJ446140 |
| 21 | 54.2 | 36.6 | 852 | 29 | CNS02108 |
| 22 | 45.8 | 30.9 | 358 | 13 | BY012221 |
| 23 | 44.8 | 30.3 | 379 | 10 | BB844013 |
| 24 | 44.8 | 30.3 | 771 | 10 | BF141725 |
| 25 | 44.8 | 30.3 | 789 | 10 | BF141725 |
| 26 | 44.8 | 30.3 | 925 | 10 | BF138310 |
| 27 | 44.8 | 30.3 | 966 | 10 | BF139214 |
| 28 | 43.2 | 29.2 | 359 | 13 | BY100690 |
| 29 | 43.2 | 29.2 | 365 | 13 | BY056132 |
| 30 | 43.2 | 29.2 | 605 | 10 | BB651860 |
| 31 | 43.2 | 29.2 | 605 | 14 | CA574608 |
| 32 | 43.2 | 29.2 | 646 | 13 | BY738045 |
| 33 | 43.2 | 29.2 | 658 | 13 | BY728644 |
| 34 | 43.2 | 29.2 | 663 | 10 | BB638923 |
| 35 | 43.2 | 29.2 | 687 | 10 | BB652985 |
| 36 | 43.2 | 29.2 | 772 | 12 | BM935138 |
| 37 | 43.2 | 29.2 | 955 | 13 | BQ934193 |
| 38 | 43.2 | 29.2 | 1802 | 11 | AK049256 |
| 39 | 43.2 | 29.2 | 2226 | 11 | AK049998 |
| 40 | 42.8 | 28.9 | 511 | 29 | CG665381 |
| 41 | 42.8 | 28.9 | 1025 | 12 | BM460629 |
| 42 | 42.4 | 28.6 | 358 | 10 | BB841045 |
| 43 | 42.4 | 28.6 | 549 | 29 | CG575900 |
| 44 | 42.2 | 28.5 | 522 | 12 | BG144705 |
| 45 | 41.6 | 28.1 | 166 | 29 | CG691593 |

ALIGNMENTS

RESULT 1
BF238155 541 bp mRNA linear EST 14-NOV-2000
LOCUS 601811880F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4054915 5',
DEFINITION mRNA sequence.
ACCESSION BF238155
VERSION BF238155.1 GI:11152074
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LNCM895 row: p column: 20
High quality sequence stop: 541.

FEATURES
source

Location/Qualifiers
1. 541
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4054915"
/issue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

ORIGIN

Query Match 100.0%; Score 148; DB 10; Length 541;
Best Local Similarity 100.0%; Pred. No. 9.7e-42;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTGATGAAACCGAGGAAGTTCTTTACCAATTCAAAATGTCGGCTGAAGG 60
DB 86 CCTCTGATGAAACCGAGGAAGTTCTTTACCAATTCAAAATGTCGGCTGAAGG 145
QY 61 TCGGCGTGAACCTACCTGCTACGTAAGAGAGCGTACAGTCAATCCTTTTC 120
DB 146 TCGGCGTGAACCTACCTGCTACGTAAGAGAGCGTACAGTCAATCCTTTTC 205
QY 121 ACTGACTTTGGTTATCTTCGCAATAG 148
DB 206 ACTGACTTTGGTTATCTTCGCAATAG 233

RESULT 2
BG686133 743 bp mRNA linear EST 01-MAY-2001
LOCUS 60263842F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4766234 5',
DEFINITION mRNA sequence.
ACCESSION BG686133
VERSION BG686133.1 GI:13917530
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 743)
NIH-MGC http://mgc.nci.nih.gov/
REFERENCE National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
AUTHORS Contact: Robert Strausberg, Ph.D.
JOURNAL Email: cgaabs-remail.nih.gov
COMMENT Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LICM1626 row: 9 column: 03
High quality sequence stop: 740.

FEATURES
source

Location/Qualifiers
1. 743
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4766234"
/issue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_48"

ORIGIN

Query Match 100.0%; Score 148; DB 12; Length 743;
Best Local Similarity 100.0%; Pred. No. 1.1e-41;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTGATGAAACCGAGGAAGTTCTTTACCAATTCAAAATGTCGGCTGAAGG 60
DB 86 CCTCTGATGAAACCGAGGAAGTTCTTTACCAATTCAAAATGTCGGCTGAAGG 145
QY 61 TCGGCGTGAACCTACCTGCTACGTAAGAGAGCGTACAGTCAATCCTTTTC 120
DB 146 TCGGCGTGAACCTACCTGCTACGTAAGAGAGCGTACAGTCAATCCTTTTC 205
QY 121 ACTGACTTTGGTTATCTTCGCAATAG 148
DB 206 ACTGACTTTGGTTATCTTCGCAATAG 233

RESULT 3
BG757089 820 bp mRNA linear EST 15-MAY-2001
LOCUS 602715124F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4855517 5',
DEFINITION mRNA sequence.
ACCESSION BG757089
VERSION BG757089.1 GI:14067742
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 820)
NIH-MGC http://mgc.nci.nih.gov/
REFERENCE National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
AUTHORS Contact: Robert Strausberg, Ph.D.
JOURNAL Email: cgaabs-remail.nih.gov
COMMENT Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LICM1704 row: 0 column: 06
High quality sequence stop: 675.

FEATURES
source

Location/Qualifiers
1. 820
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4855517"
/issue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

ORIGIN

Query Match 100.0%; Score 148; DB 12; Length 820;
Best Local Similarity 100.0%; Pred. No. 1.2e-41;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTGATGACCGGAGGAAGTTCTTTACCAATTCAAAATGTCGGCTGAGG 60
DB CCTCTGATGACCGGAGGAAGTTCTTTACCAATTCAAAATGTCGGCTGAGG 152
QY 61 TCGGCGTAGACCTACCTGCTGCTACGTAGAGAGCGGTGACAGTCTACATCTTTTC 120
DB 153 TCGGCGTAGACCTACCTGCTGCTACGTAGAGAGCGGTGACAGTCTACATCTTTTC 212
QY 121 ACTGACCTTGCTATCTTCCCAATAG 148
DB 213 ACTGACCTTGCTATCTTCCCAATAG 240

RESULT 4

LOCUS BG758510 872 bp mRNA linear EST 15-MAY-2001
DEFINITION 602712721.F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853069 5',
mRNA sequence.

ACCESSION BG758510
VERSION BG758510.1 GI:14069163
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 872)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM1698 row: i column: 06
High quality sequence stop: 836.
Location/Qualifiers

FEATURES
source

1..872
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4853069"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 148; DB 12; Length 872;
Best Local Similarity 100.0%; Pred. No. 1.2e-41;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTGATGACCGGAGGAAGTTCTTTACCAATTCAAAATGTCGGCTGAGG 60
DB 74 CCTCTGATGACCGGAGGAAGTTCTTTACCAATTCAAAATGTCGGCTGAGG 133

QY 61 TCGGCGTAGACCTACCTGCTGCTACGTAGAGAGCGGTGACAGTCTACATCTTTTC 120

DB 134 TCGGCGTAGACCTACCTGCTGCTACGTAGAGAGCGGTGACAGTCTACATCTTTTC 193

QY 121 ACTGACCTTGCTATCTTCCCAATAG 148
DB 194 ACTGACCTTGCTATCTTCCCAATAG 221

RESULT 5

LOCUS BG686876 889 bp mRNA linear EST 01-MAY-2001
DEFINITION 602650861.F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4763247 5',
mRNA sequence.

ACCESSION BG686876
VERSION BG686876.1 GI:13918273
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 889)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM1618 row: j column: 16
High quality sequence start: 6
High quality sequence stop: 727.
Location/Qualifiers

FEATURES
source

1..889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4763247"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 148; DB 12; Length 889;
Best Local Similarity 100.0%; Pred. No. 1.2e-41;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTCTGATGACCGGAGGAAGTTCTTTACCAATTCAAAATGTCGGCTGAGG 60
DB 46 CCTCTGATGACCGGAGGAAGTTCTTTACCAATTCAAAATGTCGGCTGAGG 105
QY 61 TCGGCGTAGACCTACCTGCTGCTACGTAGAGAGCGGTGACAGTCTACATCTTTTC 120
DB 106 TCGGCGTAGACCTACCTGCTGCTACGTAGAGAGCGGTGACAGTCTACATCTTTTC 165
QY 121 ACTGACCTTGCTATCTTCCCAATAG 148
DB 166 ACTGACCTTGCTATCTTCCCAATAG 193

RESULT 6
LOCUS BQ065440
DEFINITION AGENCOURT_6855061 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5829977
5', mRNA sequence.
ACCESSION BQ065440
VERSION BQ065440.1 GI:19894486
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 953)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
DNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2108 row: p column: 10
High quality sequence stop: 634.
Location/Qualifiers
1. 953
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5829977"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match 100.0%; Score 148; DB 13; Length 953;
Best Local Similarity 100.0%; Pred. No. 1.3e-41;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTCTGATGACCGGAGAGTTCTTTACCAATTCAAAATGTCGCTGGGCTAAAGG 60
DB 72 CCTCTGATGACCGGAGAGTTCTTTACCAATTCAAAATGTCGCTGGGCTAAAGG 131
QY 61 TCGGCGTGAGACCTACCTGTGCTACGTAGTGAAGAGCGGTGACAGTCTACATCCTTTTC 120
DB 132 TCGGCGTGAGACCTACCTGTGCTACGTAGTGAAGAGCGGTGACAGTCTACATCCTTTTC 191
QY 121 ACTGACTTTGGTTATCTTTCGCAATTAAG 148
DB 192 ACTGACTTTGGTTATCTTTCGCAATTAAG 219

RESULT 7
LOCUS BQ055935
DEFINITION AGENCOURT_6796291 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5808181
5', mRNA sequence.
ACCESSION BQ055935
VERSION BQ055935.1 GI:19815262

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1052)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
DNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2051 row: m column: 14
High quality sequence stop: 665.
Location/Qualifiers
1. 1052
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5808181"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match 100.0%; Score 148; DB 13; Length 1052;
Best Local Similarity 100.0%; Pred. No. 1.3e-41;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTTTGATGACCGGAGAGTTCTTTACCAATTCAAAATGTCGCTGGGCTAAAGG 60
DB 72 CCTTTGATGACCGGAGAGTTCTTTACCAATTCAAAATGTCGCTGGGCTAAAGG 131
QY 61 TCGGCGTGAGACCTACCTGTGCTACGTAGTGAAGAGCGGTGACAGTCTACATCCTTTTC 120
DB 132 TCGGCGTGAGACCTACCTGTGCTACGTAGTGAAGAGCGGTGACAGTCTACATCCTTTTC 191
QY 121 ACTGACTTTGGTTATCTTTCGCAATTAAG 148
DB 192 ACTGACTTTGGTTATCTTTCGCAATTAAG 219

RESULT 8
LOCUS AL559877
DEFINITION AL559877 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
clone CS0D00031B14 5-PRIME, mRNA sequence.
ACCESSION AL559877
VERSION AL559877.2 GI:31284008
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W. B., Gruber, C., Jessee, J., and Polyes, D.
TITLES Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)

Db 129 CCTCTGATGAAACCGAGGAAGTTCTTACCAATTGAAAATGTCGGCTGGGTAAGG 188
QY 61 TCGGCGTGAAGACCTACCTGCTCTACGTAGTAAGAGCGGTGACAGTCTACATCTTTTC 120
Db 189 TCGGCGTGAAGACCTACCTGCTCTACGTAGTAAGAGCGGTGACAGTCTACATCTTTTC 248
QY 121 ACTGACCTTGGTTATCTTGGCAATTAAG 148
Db 249 ACTGACCTTGGTTATCTTGGCAATTAAG 276

RESULT 11
Bg757392 693 bp mRNA linear EST 15-MAY-2001
LOCUS 602711022P1 NIH_MGC_48 Homo sapiens CDNA clone IMAGE:4851580 5',
DEFINITION mRNA sequence.
ACCESSION Bg757392
VERSION Bg757392
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS 1 (bases 1 to 693)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LNCM1694 row: k column: 05
High quality sequence stop: 693.
Location/Qualifiers
1..693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4851580"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

FEATURES
source
1..693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4851580"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 92.6%; Score 137; DB 12; Length 693;
Best Local Similarity 99.3%; Pred. No. 9.5e-38;
Matches 148; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CCTCTGATGAAACCGAGGAAGTTCTTACCAATTGAAAATGTCGGCTGGGTAAGG 60
Db 89 CCTCTGATGAAACCGAGGAAGTTCTTACCAATTGAAAATGTCGGCTGGGTAAGG 148
QY 61 TCGGCGTGAAGACCTACCTGCTCTACGTAGTAAGAGCGGTGACAGTCTACATCTTTT 119
Db 149 TCGGCGTGAAGACCTACCTGCTCTACGTAGTAAGAGCGGTGACAGTCTACATCTTTT 208
QY 120 CACTGACCTTGGTTATCTTGGCAATTAAG 148

Db 209 CACTGACCTTGGTTATCTTGGCAATTAAG 237

RESULT 12
BF975166 942 bp mRNA linear EST 22-JAN-2001
LOCUS 602244657P1 NIH_MGC_48 Homo sapiens CDNA clone IMAGE:4335639 5',
DEFINITION mRNA sequence.
ACCESSION BF975166
VERSION BF975166
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS 1 (bases 1 to 942)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LNCM1207 row: a column: 16
High quality sequence stop: 707.
Location/Qualifiers
1..942
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4335639"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

FEATURES

source

1..942
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4335639"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 92.6%; Score 137; DB 10; Length 942;
Best Local Similarity 99.3%; Pred. No. 1.1e-37;
Matches 148; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CCTCTGATGAAACCGAGGAAGTTCTTACCAATTGAAAATGTCGGCTGGGTAAGG 59
Db 86 CCTCTGATGAAACCGAGGAAGTTCTTACCAATTGAAAATGTCGGCTGGGTAAGG 145
QY 60 TCGGCGTGAAGACCTACCTGCTCTACGTAGTAAGAGCGGTGACAGTCTACATCTTTT 119
Db 146 TCGGCGTGAAGACCTACCTGCTCTACGTAGTAAGAGCGGTGACAGTCTACATCTTTT 205
QY 120 CACTGACCTTGGTTATCTTGGCAATTAAG 148
Db 206 CACTGACCTTGGTTATCTTGGCAATTAAG 234

RESULT 13
BX464579 853 bp mRNA linear EST 22-MAY-2003
LOCUS BX464579 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens CDNA
DEFINITION clone CS0D0003YB14 5-PRIME, mRNA sequence.
ACCESSION BX464579

VERSION BX464579.1 GI:31031641
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 853)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: secref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 6672.r For
 more information about this cluster, see
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS1DG001ZF10CP1&cluster=6672.r>. Contact :
 Feng Liang Email : fliang@life.com URL :
<http://fulllength.invitrogen.com/Invitrogen/Corporation/1600>
 Faraday Avenue Genoscope sequence ID : CS1DG001ZF10CP1.
 Location/Qualifiers
 source
 1..853
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0G003YB14"
 /tissue_type="B CELLS (RAMOS CELL LINE)"
 /cell_line="RAMOS CELL LINE"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN
 Query Match 91.9%; Score 136; DB 13; Length 853;
 Best Local Similarity 99.3%; Pred. No. 2,4e-37;
 Matches 147; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 CCTCTGATGAGACCGAGAGATTCTTTACCAATTCACAAATGTCCTGGCTAAGG 60
 DB 95 CCTCTGATGAGACCGAG-GAAGTTCTTTACCAATTCACAAATGTCCTGGCTAAGG 153
 QY 61 TCGGCGTGAAGACCTACCTGCTACGATGAGAGAGCGGACAGTCTACATCCTTTC 120
 DB 154 TCGGCGTGAAGACCTACCTGCTACGATGAGAGAGCGGACAGTCTACATCCTTTC 213
 QY 121 ACTGACTTTGTTATCTTCGCAATAG 148
 DB 214 ACTGACTTTGTTATCTTCGCAATAG 241
 RESULT 14
 BZ247020/c
 LOCUS CH230-374G16, 701 bp DNA linear GSS 12-OCT-2002
 DEFINITION CH230-374G16, genomic survey sequence.
 ACCESSION BZ247020
 VERSION BZ247020.1 GI:23907284
 KEYWORDS GSS.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 701)
 AUTHORS Zhao, S., Shetty, J., Shatsman, S., Teegaye, G., Geer, K.,
 Schwartsbeyn, A., Geirgeorgis, E., Overton, L., Russell, D., Chen, D.,
 Riggs, F., de Jong, P. and Fraser, C.M.
 TITLE Rat BAC End Sequences from Library CHORI-230 MboI segment

JOURNAL Unpublished (1999)
 COMMENT Other GSSs: CH230-374G16.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@igr.org
 Clones are derived from the rat BAC library CHORI-230
 (<http://www.chori.org/bacpac/rat230.htm>). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (<http://www.chori.org/bacpac/or> ering information.htm). BAC end
 page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
 Plate: 374 row: 6 column: 16
 Seq primer: SP6
 Class: BAC ends.
 Location/Qualifiers
 source
 1..701
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SSHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-374G16"
 /sex="female"
 /cell_type="Brain"
 /clone_id="CHORI-230 Segment 2"
 /note="Vector: PTARBAC1.3; Site 1: MboI; site 2: MboI;
 CHORI-230 Rat (BN/SSHsd/MCW) BAC library produced by
 Pieter de Jong"

ORIGIN
 Query Match 79.5%; Score 117.6; DB 28; Length 701;
 Best Local Similarity 87.2%; Pred. No. 8.5e-31;
 Matches 129; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
 QY 1 CCTTTGATGACCGAGAGATTCTTTACCAATTCACAAATGTCCTGGCTAAGG 60
 DB 602 CCTTTGATGAGAGAGAAAGATTCTTTACCACTTCACAAATGTCCTGGCTAAGG 543
 QY 61 TCGGCGTGAAGACCTACCTGCTACGATGAGAGAGCGGACAGTCTACATCCTTTC 120
 DB 542 TCGGCGTGAAGACCTACCTGCTACGATGAGAGAGCGGACAGTCTACATCCTTTC 483
 QY 121 ACTGACTTTGTTATCTTCGCAATAG 148
 DB 482 ACTGACTTTGCGCACTTCGCAATAG 455
 RESULT 15
 AJ450295
 LOCUS AJ450295 623 bp mRNA linear EST 19-APR-2002
 DEFINITION AJ450295 riken1 Gallus gallus cDNA clone 24m21r1, mRNA sequence.
 ACCESSION AJ450295
 VERSION AJ450295.1 GI:20217516
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 623)
 AUTHORS Buerstedde, J.M.
 TITLE Gallus gallus bursal lymphocyte EST
 JOURNAL Unpublished (2002)
 COMMENT Contact: Buerstedde JM
 Cellular Immunology
 Heinrich-Pette-Institute
 Martinistr. 52, 20251 Hamburg, Germany
 Email: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.
 Location/Qualifiers
 source
 1..623

```
/organism="Gallus gallus"  
/mol_type="mRNA"  
/db_xref="taxon:9031"  
/clone="2421x1"  
/cell_type="bursal lymphocyte"  
/dev_stage="2-3 weeks old"  
/clone_id="r1ken1"  
/note="CB inbred strain"
```

ORIGIN

```
Query Match      63.2%; Score 93.6; DB 9; Length 623;  
Best Local Similarity 77.0%; Pred. No. 3.2e-22;  
Matches 114; Conservative 0; Mismatches 34; Indels 0; Gaps 0;  
  
QY      1 CCTCTGATGAACCGAGGAAAGTTCTTACCAATCAAAAATGCCGCTGAGGTAAGG 60  
        |||||  
Db      165 CCTCTGATGAAGAGAGAGCTCTTCTCTACAATTTCAGAACTGCCGCTGAGGTAAGG 224  
  
QY      61 TCGGCGTGAAGCCTTCTGTGCTACGTAGTGAAGGCGGTGAACGTGCTACATCCTTTTC 120  
        |||||  
Db      225 CCGTCGTGAACCTACCTCTGTATGTGGAAGCGCGGTGAACGTGCTACATCATGCTC 284  
  
QY      121 ACTGACTTTGTTATCTTGCAATAG 148  
        |||||  
Db      285 CCTGACTTTGATACCTGCGTAACAG 312
```

Search completed: March 13, 2004, 00:39:02
Job time : 474.394 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 12:27:59 ; Search time 853.181 Seconds

(without alignments)
13767.261 Million cell updates/sec

Title: US-09-966-880A-13

Perfect score: 271

Sequence: 1 aagcgctgcacgttggaatt.....agccatcatgaccttcaag 271

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_hcg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_hcg_hum:*

31: em_hcg_inv:*

32: em_hcg_other:*

33: em_hcg_mus:*

34: em_hcg_pln:*

35: em_hcg_rtd:*

36: em_hcg_man:*

37: em_hcg_vtl:*

38: em_sy:*

39: em_higo_hum:*

40: em_higo_mus:*

41: em_higo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 271 | 100.0 | 271 | 6 BD016838 | BD016838 Novel cyt |
| 2 | 271 | 100.0 | 547 | 12 AF529856 | AF529856 Cricetulu |
| 3 | 271 | 100.0 | 591 | 12 AF529855 | AF529855 Cricetulu |
| 4 | 271 | 100.0 | 595 | 12 AF529834 | AF529834 Mus muscu |
| 5 | 271 | 100.0 | 596 | 12 AF529815 | AF529815 Homo sapi |
| 6 | 271 | 100.0 | 596 | 9 AF529822 | AF529822 Homo sapi |
| 7 | 271 | 100.0 | 596 | 9 AF529823 | AF529823 Homo sapi |
| 8 | 271 | 100.0 | 596 | 9 AF529825 | AF529825 Homo sapi |
| 9 | 271 | 100.0 | 596 | 12 AF529829 | AF529829 Mus muscu |
| 10 | 271 | 100.0 | 596 | 12 AF529831 | AF529831 Mus muscu |
| 11 | 271 | 100.0 | 596 | 12 AF529832 | AF529832 Mus muscu |
| 12 | 271 | 100.0 | 596 | 12 AF529833 | AF529833 Mus muscu |
| 13 | 271 | 100.0 | 596 | 12 AF529835 | AF529835 Mus muscu |
| 14 | 271 | 100.0 | 596 | 12 AF529836 | AF529836 Mus muscu |
| 15 | 271 | 100.0 | 597 | 9 BT007402 | BT007402 Homo sapi |
| 16 | 271 | 100.0 | 597 | 12 AF529842 | AF529842 Cricetulu |
| 17 | 271 | 100.0 | 597 | 12 AF529843 | AF529843 Cricetulu |
| 18 | 271 | 100.0 | 597 | 12 AF529846 | AF529846 Cricetulu |
| 19 | 271 | 100.0 | 597 | 12 AF529847 | AF529847 Cricetulu |
| 20 | 271 | 100.0 | 597 | 12 AF529851 | AF529851 Cricetulu |
| 21 | 271 | 100.0 | 597 | 12 AF529853 | AF529853 Cricetulu |
| 22 | 271 | 100.0 | 597 | 12 BT008226 | BT008226 Synthetic |
| 23 | 271 | 100.0 | 1828 | 9 BC006296 | BC006296 Homo sapi |
| 24 | 271 | 100.0 | 2791 | 9 AB040431 | AB040431 Homo sapi |
| 25 | 271 | 100.0 | 2818 | 6 BD016833 | BD016833 Novel cyt |
| 26 | 271 | 100.0 | 6564 | 6 BD016835 | BD016835 Novel cyt |
| 27 | 271 | 100.0 | 11204 | 6 BD016860 | BD016860 Novel cyt |
| 28 | 271 | 100.0 | 11204 | 9 AB040430 | AB040430 Homo sapi |
| 29 | 271 | 100.0 | 71132 | 9 AC092184 | AC092184 Homo sapi |
| 30 | 269.4 | 99.4 | 577 | 12 AF529840 | AF529840 Mus muscu |
| 31 | 269.4 | 99.4 | 596 | 9 AF529816 | AF529816 Homo sapi |
| 32 | 269.4 | 99.4 | 596 | 9 AF529817 | AF529817 Homo sapi |
| 33 | 269.4 | 99.4 | 596 | 9 AF529818 | AF529818 Homo sapi |
| 34 | 269.4 | 99.4 | 596 | 9 AF529819 | AF529819 Homo sapi |
| 35 | 269.4 | 99.4 | 596 | 9 AF529820 | AF529820 Homo sapi |
| 36 | 269.4 | 99.4 | 596 | 9 AF529821 | AF529821 Homo sapi |
| 37 | 269.4 | 99.4 | 596 | 9 AF529824 | AF529824 Homo sapi |
| 38 | 269.4 | 99.4 | 596 | 9 AF529826 | AF529826 Homo sapi |
| 39 | 269.4 | 99.4 | 596 | 9 AF529827 | AF529827 Homo sapi |
| 40 | 269.4 | 99.4 | 596 | 12 AF529830 | AF529830 Mus muscu |
| 41 | 269.4 | 99.4 | 596 | 12 AF529837 | AF529837 Mus muscu |
| 42 | 269.4 | 99.4 | 596 | 12 AF529839 | AF529839 Mus muscu |
| 43 | 269.4 | 99.4 | 597 | 12 AF529828 | AF529828 Mus muscu |
| 44 | 269.4 | 99.4 | 597 | 12 AF529841 | AF529841 Cricetulu |
| 45 | 269.4 | 99.4 | 597 | 12 AF529844 | AF529844 Cricetulu |

ALIGNMENTS

RESULT 1

BD016838

LOCUS BD016838

DEFINITION Novel cytidine deaminase.

ACCESSION BD016838

VERSION BD016838.1 GI:22558014

KEYWORDS UP 2001245669-A/11.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 271)

AUTHORS Honjo, T. and Muramatsu, M.

TITLE Novel cytidine deaminase

JOURNAL Patent: JP 2001245669-A 11 11-SEP-2001;

| COMMENT | | | | | | | |
|--------------------------------|--|-------------|----|--|--|--|--|
| JAPAN TOBACCO INC,TASUKU HONJO | | | | | | | |
| OS | Homo sapiens (human) | | | | | | |
| PX | JR 2001245659-A/11 | | | | | | |
| PD | 11-SEP-2001 | | | | | | |
| PF | 28-MAR-2000 JP 2000092981 | | | | | | |
| PI | TASUKU HONJO,MASAMICHI MORAYATSU | | | | | | |
| PC | C12N5/09,A6IKX39/395,A6IKX39/395,A6IP1/00,A6IP11/06,A6IP13/12, | | | | | | |
| PC | A6IP17/00, | | | | | | |
| PC | A6IP27/02,A6IP27/16,A6IP37/02,A6IP37/08,COTK16/18,C12N1/19, PC | | | | | | |
| | C12N1/21, | | | | | | |
| | PC C12N5/10,C12N9/78,C12P21/02,C12P21/08/(C12N1/21,C12R1:19), PC | | | | | | |
| | (C12M5/10,C12R1:91),C12N15/00,C12N5/00,(C12N5/00,C12R1:91) CC | | | | | | |
| FH | Key Location/Qualifiers. | | | | | | |
| | | | | | | | |
| FEATURES | | | | | | | |
| source | 1..271 | | | | | | |
| | /organism="Homo sapiens" | | | | | | |
| | /mol_type="genomic DNA" | | | | | | |
| | /db_xref="taxon:9606" | | | | | | |
| ORIGIN | | | | | | | |
| Query Match | 100.0%; Score 271; DB 6; | Length 271; | | | | | |
| Best Local Similarity | 100.0%; Pred. No. 6.4e-50; | | | | | | |
| Matches 271; Conservative | 0; Mismatches 0; Indels | 0; Gaps | 0; | | | | |
| OY | 1 AACGGCTGCGACGTGAATTTTCCTCCGCCGACTCATCTGGAGTGGAACTGACCT 60 | | | | | | |
| Dd | 1 AAGGGCTGCAGTGGAATTGCTCTTCCTCCGCAATCTCGAGTGGAGACCTAGAACCCT 60 | | | | | | |
| OY | 61 GGCCCGCTGTCACCGGGTCACTGTTACAACCTCCTGAGACCCCCTGTAAGACTGCCCCGA 120 | | | | | | |
| Dd | 61 GGCCCGCTGTCACCGGGTCACTGTTACAACCTCCTGAGACCCCCTGTAAGACTGCCCCGA 120 | | | | | | |
| OY | 121 CATGTGGCCGACTTTCTGCGAAGGAACCCCAAACCTCACTTGAGATCTTCAACCGCGGC 180 | | | | | | |
| Dd | 121 CATGTGGCCGACTTTCTGCGAAGGAACCCCAAACCTCACTTGAGATCTTCAACCGCGGC 180 | | | | | | |
| OY | 181 CTCCTACTTCTGTGAGAGAACCGGAGAGCTGAGCCCCGAGGGGCTGGACCGCGCC 240 | | | | | | |
| Dd | 181 CTCCTACTTCTGTGAGAGAACCGGAGAGCTGAGCCCCGAGGGGCTGGACCGCGCC 240 | | | | | | |
| OY | 241 GGGGTGCAAAATGACCATCATGACCTTCAAG 271 | | | | | | |
| Dd | 241 GGGGTGCAAAATGACCATCATGACCTTCAAG 271 | | | | | | |
| <hr/> | | | | | | | |
| RESULT 2 | | | | | | | |
| AF529856 | 547 bp mRNA linear SYN 21-AUG-2003 | | | | | | |
| LOCUS | Cricetus griseus transgenic clone 16 Homo Sapiens AID (AID) mRNA, | | | | | | |
| DEFINITION | complete cds. | | | | | | |
| ACCESSION | AF529856 | | | | | | |
| VERSION | AF529856.1 GI:22297299 | | | | | | |
| KEYWORDS | . | | | | | | |
| SOURCE | Cricetus griseus (Chinese hamster) | | | | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Cricetus. | | | | | | |
| REFERENCE | 1 (bases 1 to 547) Martin,A. and Schaffr,M.D. Somatic hypermutation of the AID transgene in B and non-B cells Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002) | | | | | | |
| AUTHORS | Martin,A. and Schaffr,M.D. | | | | | | |
| JOURNAL | Direct Submission | | | | | | |
| MEDLINE | 22222991 | | | | | | |
| PubMed | 12202747 | | | | | | |
| REFERENCE | 2 (bases 1 to 547) Martin,A. and Schaffr,M.D. Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA | | | | | | |
| AUTHORS | location/Qualifiers | | | | | | |
| TITLE | 1..547 | | | | | | |
| JOURNAL | /organism="Cricetus griseus" | | | | | | |
| FEATURES | /mol_type="mRNA" | | | | | | |
| Source | | | | | | | |

| | | | | | | | | | | |
|---------------------------|--|--|-----------|-------------|--|--|--|--|--|--|
| source | /db_xref="taxon:10029" /clone="16" /cell_line="CHO" /transgenic 1..547 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" <1..547 /gene="AID" /gene="AID" <1..547 /gene="AID" /codon_start=2 /product="AID" /protein_id="AA05443.1" /db_xref="GI:2287300" /translation="VRNAKRREITLYCYVVRKDSATSPSLDPGLYLNKNGCHVELLP LRYSIDMDLPGRCYRTATWTFSPSPCYDCARHVADFLRGNPNLSLRFTALVPCEDR KAEBEGRLIRPAGVOALMTFMDYFCNMTFVENHRTFKAMEGLHENSRLSRQLR RIILPLVEVDINDAPRTWR" | | | | | | | | | |
| gene | | | | | | | | | | |
| CDS | | | | | | | | | | |
| ORIGIN | | | | | | | | | | |
| Query Match | 100.0%; | Score 271; | DB 12; | Length 547; | | | | | | |
| Best Local Similarity | 100.0%; | Pred. No. 5.9e-50; | | | | | | | | |
| Matches 271; Conservative | 0; | Mismatches 0; | Indels 0; | Gaps 0; | | | | | | |
| QY | 1 | AACGGCTGCCACGTGGAAATTCCTTCTTCCTCGCTACATCTTGGACTGGAGACTTGAGACCT | 60 | | | | | | | |
| Db | 107 | AACGGCTGCCACGTGGAAATTCCTTCTTCCTCGCTACATCTTGGACTGGAGACTTGAGACCT | 166 | | | | | | | |
| QY | 61 | GCGCGCTGCTACGGGGCACTGGTTCACCTCCCTGGAGCCCTGCTAGACTGAGCCGA | 120 | | | | | | | |
| Db | 167 | GCGCGCTGCTACGGGGCACTGGTTCACCTCCCTGGAGCCCTGCTAGACTGAGCCGA | 226 | | | | | | | |
| QY | 121 | CATGTGGCCGACTTTCGCGAGGAGAACCCCACTCAGTCTGAGGATCTTTCACCGCGGC | 180 | | | | | | | |
| Db | 227 | CATGTGGCCGACTTTCGCGAGGAGAACCCCACTCAGTCTGAGGATCTTTCACCGCGGC | 286 | | | | | | | |
| QY | 181 | CTTACTTCTGTGAGAACCGGACGCTGAGGCGGAGGCGGCTGCGCGGCTGCAACCGGCC | 240 | | | | | | | |
| Db | 287 | CTTACTTCTGTGAGAACCGGACGCTGAGGCGGAGGCGGCTGCGCGGCTGCAACCGGCC | 346 | | | | | | | |
| QY | 241 | GGGGTGCAATAGCCATCATGACCTTCAAG | 271 | | | | | | | |
| Db | 347 | GGGGTGCAATAGCCATCATGACCTTCAAG | 377 | | | | | | | |
| RESULT 3 | | | | | | | | | | |
| AF529855 | | | | | | | | | | |
| LOCUS | 591 bp, mRNA, linear, SYN 21-AUG-2003 | | | | | | | | | |
| DEFINITION | Cricetulus griseus transgenic clone 15 Homo sapiens AID (AID) mRNA, complete cds. | | | | | | | | | |
| ACCESSION | AF529855 | | | | | | | | | |
| VERSION | AF529855.1 GI:22297297 | | | | | | | | | |
| KEYWORDS | | | | | | | | | | |
| SOURCE | | | | | | | | | | |
| ORGANISM | Cricetulus griseus (Chinese hamster) | | | | | | | | | |
| REFERENCE | Cricetulus griseus | | | | | | | | | |
| AUTHORS | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Cricetulus. | | | | | | | | | |
| TITLE | 1 (bases 1 to 591) | | | | | | | | | |
| JOURNAL | Martin, A. and Scharff, M.D. | | | | | | | | | |
| MEDLINE | Somatic hypermutation of the AID transgene in B and non-B cells | | | | | | | | | |
| PubMed | 22322391 | | | | | | | | | |
| REFERENCE | 2 (bases 1 to 591) | | | | | | | | | |
| AUTHORS | Martin, A. and Scharff, M.D. | | | | | | | | | |
| TITLE | Direct Submission | | | | | | | | | |
| JOURNAL | Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Ave, Chain 404, Bronx, NY 10461, USA | | | | | | | | | |
| FEATURES | Location/Qualifiers | | | | | | | | | |
| source | 1..591 | | | | | | | | | |


```

gene
  clone="Ramos 1"
  .>596
  /gene="AID"
  1..>596
  /gene="AID"
  /note="Integrated into Burkitt's lymphoma cell line Ramos"
  /codon_start=1
  /product="AID"
  /protein_id="AAM95402.1"
  /db_xref="GI:22297218"
  /translation="MDSILMMNRKFLYQFNKRWAKGRRETYLCTVYKRSDATSFSLL
DFFGLRNKNGCHEVELFLRYISDWDLPGRGCRYTWFTFSWSPCYDCARHVDLFRGNP
NLSRIFFARLYPCEDRRAEPEGRLRLRAGVQALIMTFKDYFCMNTFVENHRTFRK
AMEGLHENSVALSRQIRRLILPLVEYVDLDRAPRTLG."

ORIGIN
Query Match 100.0%; Score 271; DB 9; Length 596;
Best Local Similarity 100.0%; Pred. No. 5,9e-50;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGCTGCCACGTGGAATTGCTCTTCTCCGCTACATCTGGAGTGGACCTTGACCTT 60
Dh 157 AACGGCTGCCACGTGGAATTGCTCTTCTCCGCTACATCTGGAGTGGACCTT 216
QY 61 GGCGGCTGCTACCGCGTACCGTGGTACCTCCTGGAGCCCGCTGTACAGCTGTGCCGGA 120
Dh 217 GGCGGCTGCTACCGCGTACCGTGGTACCTCCTGGAGCCCGCTGTACAGCTGTGCCGGA 276
QY 121 CATGTGGCGCACTTCTCTCGAGAGGAACCCCAACCTCACTGTAGAGATCTTACCGCGGC 180
Dh 277 CATGTGGCGCACTTCTCTCGAGAGGAACCCCAACCTCACTGTAGAGATCTTACCGCGGC 336
QY 181 CTCTACTTCTGTGAGAGACCGGCAAGGCTGAGCCCGAGGGGCTGCGCGGCTGCACCGCGCC 240
Dh 337 CTCTACTTCTGTGAGAGACCGGCAAGGCTGAGCCCGAGGGGCTGCGCGGCTGCACCGCGCC 396
QY 241 GGGGTGCAATATGACCATATGATGACCTTCAAG 271
Dh 397 GGGGTGCAATATGACCATATGATGACCTTCAAG 427

RESULT 6
AF529822 596 bp mRNA linear PRI 17-SEP-2002
LOCUS AF529822 Homo sapiens clone Ramos 8 AID (AID) mRNA, partial cds.
DEFINITION Homo sapiens clone Ramos 8 AID (AID) mRNA, partial cds.
ACCESSION AF529822
VERSION AF529822.1 GI:22297231
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 596)
Martin,A. and Scharff,M.D.
Somatic hypermutation of the AID transgene in B and non-B cells
Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
2 (bases 1 to 596)
Martin,A. and Scharff,M.D.
Direct Submission
Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA
location/Qualifiers
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 8"
1..>596
/gene="AID"
1..>596
/gene="AID"
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1

```

| | | | | |
|---------------------------|--|--------------------|-----------|-----------------|
| ORIGIN | /product="AID" /protein_id="AAM95409.1" /db_xref="GI:22297232" /translation="MDSLLMKRRKFLYQFNVMFAKARETYLCTVYKRDRTSFL DFGLRNKNGCHVELFLRYISDMDDPGRCYVTWFTSMPCYDCAHYADLRNP NLSIRITARLYFCEDRKABEGRLRIHRAVGQIATMTFDYFCWNTFVENHERTEK AMEGIHENSVRLSRQLRRILLPLYEVDLDADAPRTGL" | | | |
| Query Match | 100.0%; | Score 271; | DB 9; | Length 596; |
| Best Local Similarity | 100.0%; | Pred. No. 5,9e-50; | | |
| Matches 271, Conservative | 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| OY | 1 AACGGCTGCCACCGGGAATGCTCTCTCTCGCGACATCTCGACCTGGACCTTAGACCT 60 DB 157 AACGGCTGCCACCGGGAATGCTCTCTCTCGCGACATCTCGACCTGGACCTTAGACCT 216 61 GGCGCGTGTACCGCGCTACCGTGGTTACCTCTCGAGCCCTGTGCTAGCTGGACCGGA 120 217 GGCGCGTGTACCGCGCTACCGTGGTTACCTCTCGAGCCCTGTGCTAGCTGGACCGGA 276 OY 121 CATGTGCGCGCATTTCTTGCGAGAGGACCCCAACCTCACTGCTGAGATCTTACCGCGCGC 180 DB 277 CATGTGCGCGCATTTCTTGCGAGAGGACCCCAACCTCACTGCTGAGATCTTACCGCGCGC 336 OY 181 CTGACTTCTGTGAGGACCGCGAGGCTGAGCCCGGAGGGGCTGGCGCGGCTGCACCGCGCC 240 DB 337 CTGACTTCTGTGAGGACCGCGAGGCTGAGCCCGGAGGGGCTGGCGCGGCTGCACCGCGCC 396 OY 241 GGCGTGCAATATGATCATGATGATCTTCAAG 271 DB 397 GGCGTGCAATATGATCATGATGATCTTCAAG 427 | | | |
| RESULT 7 | | | | |
| AF529823 | 596 bp | mRNA | linear | PRI 17-SEP-2002 |
| LOCUS | | | | |
| DEFINITION | Homo sapiens clone Ramos 9 AID (AID) mRNA, partial cds. | | | |
| ACCESSION | AF529823 | | | |
| VERSION | AF529823.1 GI:22297233 | | | |
| KEYWORDS | | | | |
| SOURCE | Homo sapiens (human) | | | |
| ORGANISM | Homo sapiens (human) | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 596) 2 (bases 1 to 596) Martin,A. and Scharff,M.D. Somatic hypermutation of the AID transgene in B and non-B cells Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002) | | | |
| AUTHORS | Martin,A. and Scharff,M.D. | | | |
| JOURNAL | Direct Submission | | | |
| TITLE | Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA | | | |
| JOURNAL | Location/Qualifiers | | | |
| FEATURES | 1..596 | | | |
| source | /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="Ramos 9" 1..596 /gene="AID" 1..596 /gene="AID" /note="Integrated into Burkitt's lymphoma cell line Ramos" /codon_start=1 /product="AID" /protein_id="AAM95410.1" /db_xref="GI:22297234" /translation="MDSLLMKRRKFLYQFNVMFAKARETYLCTVYKRDRTSFL DFGLRNKNGCHVELFLRYISDMDDPGRCYVTWFTSMPCYDCAHYADLRNP NLSIRITARLYFCEDRKABEGRLRIHRAVGQIATMTFDYFCWNTFVENHERTEK AMEGIHENSVRLSRQLRRILLPLYEVDLDADAPRTGL" | | | |
| gene | | | | |
| CDS | | | | |

ORIGIN

Query Match 100.0%; Score 271; DB 9; Length 596;
Best Local Similarity 100.0%; Pred. No. 5.9e-50;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGCTGCCACGTGAATTGCTCTTCCTCCGCTACATCTCGAGCTGGAACCTAGACCT 60
DB 157 AACGGCTGCCACGTGAATTGCTCTTCCTCCGCTACATCTCGAGCTGGAACCTAGACCT 216
QY 61 GGGCGCTGCTACCGCGCTACCTGTTTCACTCTGAGAGCCCTGCTACGACTGTGCCGA 120
DB 217 GGGCGCTGCTACCGCGCTACCTGTTTCACTCTGAGAGCCCTGCTACGACTGTGCCGA 276
QY 121 CATGTGCGCAGCTTTCTGCGAGGGAACCCCACTGATCTGAGATCTTCAACCGCGGC 180
DB 277 CATGTGCGCAGCTTTCTGCGAGGGAACCCCACTGATCTGAGATCTTCAACCGCGGC 336
QY 181 CTCTACTTCTGTGAGGACCGCAAGCTGAGCCCGAGGCGCTGCGCGCTGACCGCGCC 240
DB 337 CTCTACTTCTGTGAGGACCGCAAGCTGAGCCCGAGGCGCTGCGCGCTGACCGCGCC 396
QY 241 GGGGTGCAAAATGACCATCATGACCTTCAAG 271
DB 397 GGGGTGCAAAATGACCATCATGACCTTCAAG 427

RESULT 8
AF529825 596 bp mRNA linear PRI 17-SEP-2002

LOCUS Homo sapiens clone Ramos 11 AID (AID) mRNA, partial cds.
DEFINITION AF529825
ACCESSION AF529825.1 GI:22297237

VERSION AF529825.1 GI:22297237
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 596)
AUTHORS Martin,A. and Scharff,M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)

REFERENCE 2 (bases 1 to 596)
AUTHORS Martin,A. and Scharff,M.D.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA

FEATURES
location/Qualifiers
1..596
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 11"
1..>596
/gene="AID"
1..>596
/gene="AID"
/note="Integrated into Burkitt's lymphoma cell line Ramos"

CDS
1..>596
/gene="AID"
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="AA095412.1"
/db_xref="GI:22297238"
/translation="MDSLMNRKFLYQPKVNRWAKGRREAVLCYVVRDSTFSL
DFGLRNKGCHVEILFLRYISDMIDPGRCHVTFMTSPCYDCARHVAADFLRGNP
NLSIRITPAIVCEBRKAPBGLRLHRAVOIALMTFKDYFCWNTFVENHERTEK
AMEGLHNSVRLSRQLRLILPLIYEVDDLRDAFRTIGL"

ORIGIN

Query Match 100.0%; Score 271; DB 9; Length 596;
Best Local Similarity 100.0%; Pred. No. 5.9e-50;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGCTGCCACGTGAATTGCTCTTCCTCCGCTACATCTCGAGCTGGAACCTAGACCT 60

DB 157 AACGGCTGCCACGTGAATTGCTCTTCCTCCGCTACATCTCGAGCTGGAACCTAGACCT 216

QY 61 GGGCGCTGCTACCGCGCTACCTGTTTCACTCTGAGAGCCCTGCTACGACTGTGCCGA 120

DB 217 GGGCGCTGCTACCGCGCTACCTGTTTCACTCTGAGAGCCCTGCTACGACTGTGCCGA 276

QY 121 CATGTGCGCAGCTTTCTGCGAGGGAACCCCACTGATCTGAGATCTTCAACCGCGGC 180

DB 277 CATGTGCGCAGCTTTCTGCGAGGGAACCCCACTGATCTGAGATCTTCAACCGCGGC 336

QY 181 CTCTACTTCTGTGAGGACCGCAAGCTGAGCCCGAGGCGCTGCGCGCTGACCGCGCC 240

DB 337 CTCTACTTCTGTGAGGACCGCAAGCTGAGCCCGAGGCGCTGCGCGCTGACCGCGCC 396

QY 241 GGGGTGCAAAATGACCATCATGACCTTCAAG 271

DB 397 GGGGTGCAAAATGACCATCATGACCTTCAAG 427

RESULT 9
AF529829 596 bp mRNA linear SYN 17-SEP-2002

LOCUS Mus musculus clone 2 transgenic Homo sapiens AID (AID) mRNA,
complete cds.
DEFINITION AF529829
ACCESSION AF529829.1 GI:22297245

VERSION AF529829.1 GI:22297245
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 596)
AUTHORS Martin,A. and Scharff,M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)

REFERENCE 2 (bases 1 to 596)
AUTHORS Martin,A. and Scharff,M.D.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA

FEATURES
location/Qualifiers
1..596
source
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="2"
/cell_type="hybridoma P1-5"
/transgenic
1..596
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..>596
/gene="AID"
1..>596
/gene="AID"
/note="Integrated into Burkitt's lymphoma cell line Ramos"

CDS
1..>596
/gene="AID"
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="AA095416.1"
/db_xref="GI:22297246"
/translation="MDRLMNRKFLYQPKVNRWAKGRREAVLCYVVRDSTFSL
DFGLRNKGCHVEILFLRYISDMIDPGRCHVTFMTSPCYDCARHVAADFLRGNP
NLSIRITPAIVCEBRKAPBGLRLHRAVOIALMTFKDYFCWNTFVENHERTEK
AMEGLHNSVRLSRQLRLILPLIYEVDDLRDAFRTIGL"

ORIGIN

Query Match 100.0%; Score 271; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 5.9e-50;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGCTGCCACGTGAATTGCTCTTCCTCCGCTACATCTCGAGCTGGAACCTAGACCT 60

Db 157 AACGCGTCCACGCTGGAATTGCTCTCTCCGCTACATCTCGGACTGGAGCTAGACCTT 216
QY 61 GGCCGCTGTACACGCGCTCACTGATTTCACCTCTGTGAGCCCTGTACGACTGTGCCCCA 120
Db 217 GGCCGCTGTACACGCGCTCACTGATTTCACCTCTGTGAGCCCTGTACGACTGTGCCCCA 276
QY 121 CATGTGGCCGACCTTTCTGCGAGGAAACCCCAACCTCAGTCTGAGATCTTCAACCGCGCC 180
Db 277 CATGTGGCCGACCTTTCTGCGAGGAAACCCCAACCTCAGTCTGAGATCTTCAACCGCGCC 336
QY 181 CTCTACTTCTGTAGAGACCCGCAAGGCTGAGCCGAGGGGCTGGGGGCTGCAACCGCGCC 240
Db 337 CTCTACTTCTGTAGAGACCCGCAAGGCTGAGCCGAGGGGCTGGGGGCTGCAACCGCGCC 396
QY 241 GGGGTGCAATAGGCATCATGACCTTCAAG 271
Db 397 GGGGTGCAATAGGCATCATGACCTTCAAG 427

RESULT 10
AF529831 596 bp mRNA linear SYN 17-SEP-2002
LOCUS Mus musculus clone 4 transgenic Homo sapiens AID (AID) mRNA,
DEFINITION complete cds.
ACCESSION AF529831
VERSION AF529831.1 GI:22297249
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 596)
TITLE Martin, A. and Scharff, M.D.
JOURNAL Somatic hypermutation of the AID transgene in B and non-B cells
AUTHORS Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
2 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA
FEATURES
source
1..596
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="5"
/cell_type="hybridoma P1-5"
/transgenic
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..>596
/gene="AID"
1..>596
/gene="AID"
/codon_start=1
/product="AID"
/protein_id="AA95419.1"
/db_xref="GI:22297250"
/translation="MDSLMMRRKFLYQFNKVMKGRRETYLCYVVKRDSATSPSL
DFGLRNKNGCHVELFLRYSIDWDLDPRGCRVYMTFTSWSPCYDCARHVADELKGNP
NLSIRFTALVFCEDRKAPBGLRLHRAVGQIALMTFKDYFCMNTFVENHRTFK
AMEGLHNSVRLSRQFRRLILPLYEVDLDLDAFRTLGL"

ORIGIN

Query Match 100.0%; Score 271; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 5.9e-50;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACGCGTCCACGCTGGAATTGCTCTCTCCGCTACATCTCGGACTGGAGCTAGACCTT 60
Db 157 AACGCGTCCACGCTGGAATTGCTCTCTCCGCTACATCTCGGACTGGAGCTAGACCTT 216

QY 61 GGCCGCTGTACACGCGCTCACTGATTTCACCTCTGTGAGCCCTGTACGACTGTGCCCCA 120
Db 217 GGCCGCTGTACACGCGCTCACTGATTTCACCTCTGTGAGCCCTGTACGACTGTGCCCCA 276
QY 121 CATGTGGCCGACCTTTCTGCGAGGAAACCCCAACCTCAGTCTGAGATCTTCAACCGCGCC 180
Db 277 CATGTGGCCGACCTTTCTGCGAGGAAACCCCAACCTCAGTCTGAGATCTTCAACCGCGCC 336
QY 181 CTCTACTTCTGTAGAGACCCGCAAGGCTGAGCCGAGGGGCTGGGGGCTGCAACCGCGCC 240
Db 337 CTCTACTTCTGTAGAGACCCGCAAGGCTGAGCCGAGGGGCTGGGGGCTGCAACCGCGCC 396
QY 241 GGGGTGCAATAGGCATCATGACCTTCAAG 271
Db 397 GGGGTGCAATAGGCATCATGACCTTCAAG 427

RESULT 11
AF529832 596 bp mRNA linear SYN 17-SEP-2002
LOCUS Mus musculus clone 5 transgenic Homo sapiens AID (AID) mRNA,
DEFINITION complete cds.
ACCESSION AF529832
VERSION AF529832.1 GI:22297251
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 596)
TITLE Martin, A. and Scharff, M.D.
JOURNAL Somatic hypermutation of the AID transgene in B and non-B cells
AUTHORS Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
2 (bases 1 to 596)
AUTHORS Martin, A. and Scharff, M.D.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA
FEATURES
source
1..596
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="5"
/cell_type="hybridoma P1-5"
/transgenic
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..>596
/gene="AID"
1..>596
/gene="AID"
/codon_start=1
/product="AID"
/protein_id="AA95419.1"
/db_xref="GI:22297252"
/translation="MDSLMMRRKFLYQFNKVMKGRRETYLCYVVKRDSATSPSL
DFGLRNKNGCHVELFLRYSIDWDLDPRGCRVYMTFTSWSPCYDCARHVADELKGNP
NLSIRFTALVFCEDRKAPBGLRLHRAVGQIALMTFKDYFCMNTFVENHRTFK
AMEGLHNSVRLSRQFRRLILPLYEVDLDLDAFRTLGL"

ORIGIN

Query Match 100.0%; Score 271; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 5.9e-50;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACGCGTCCACGCTGGAATTGCTCTCTCCGCTACATCTCGGACTGGAGCTAGACCTT 60
Db 157 AACGCGTCCACGCTGGAATTGCTCTCTCCGCTACATCTCGGACTGGAGCTAGACCTT 216

QY 61 GGCCGCTGTACCGCGTCACTGATTCACCTCTGTGAGCCCTGTGACGACTGTGCCCCGA 120
DB 217 GGCCGCTGTACCGCGTCACTGATTCACCTCTGTGAGCCCTGTGACGACTGTGCCCCGA 276
QY 121 CATGTGGCGGACCTTTCTGCGAGGGAACCCCAACCTCAGTGTGAGATCTTTCACCGGCGC 180
DB 277 CATGTGGCGGACCTTTCTGCGAGGGAACCCCAACCTCAGTGTGAGATCTTTCACCGGCGC 336
QY 181 CTCTACTTCTGTGAGGACCGCAAGGCTGAGCCCGAGGCGCTGCGGCGCTGACCGCGCC 240
DB 337 CTCTACTTCTGTGAGGACCGCAAGGCTGAGCCCGAGGCGCTGCGGCGCTGACCGCGCC 396
QY 241 GGCGTGCAATAGCCATCATGACCTTCAAG 271
DB 397 GGCGTGCAATAGCCATCATGACCTTCAAG 427

RESULT 12
AF529833 596 bp mRNA linear SYN 17-SEP-2002
LOCUS Mus musculus clone 6 transgenic Homo sapiens AID (AID) mRNA,
DEFINITION complete cds.
ACCESSION AF529833
VERSION AF529833.1 GI:22297253
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE Martin, A. and Scharff, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
AUTHORS 2 (bases 1 to 596)
MARTIN, A. and SCHARFF, M.D.
TITLE Direct Submision
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA
FEATURES
source
1..596
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="6"
/cell_type="hybridoma PI-5"
/transgenic
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..>596
/gene="AID"
1..>596
/gene="AID"
/codon_start=1
/product="AID"
/protein_id="AA095420.1"
/db_xref="GI:22297254"
/translation="MDSLMNRKFLYQKRVNMAKGRRETYVYVYKRRDSATSFSL
DFGLRNKNGCHELLFLRYISMDLDPGRCYAVTWFTMSPCYDCAHVADFLRGNP
NLSIRITFALVYFCEDRKAPEGRLRHAGVOIALMTFKDYFCWNTFVENHRTFK
AMEGLHNSVRLSRQLRILLPLYEVDLDRARFTLGL"

ORIGIN
Query Match 100.0%; Score 271; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 5.9e-50;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCTGCACGCGAATTGCTTCTCGCTACATCTCGGACTGGACCTAGACCT 60
DB 157 AAGCGCTGCACGCGAATTGCTTCTCGCTACATCTCGGACTGGACCTAGACCT 216
QY 61 GGCGCTGTACCGCGTCACTGATTCACCTCTGTGAGCCCTGTGACGACTGTGCCCCGA 120

DB 217 GGCCGCTGTACCGCGTCACTGATTCACCTCTGTGAGCCCTGTGACGACTGTGCCCCGA 276
QY 121 CATGTGGCGGACCTTTCTGCGAGGGAACCCCAACCTCAGTGTGAGATCTTTCACCGGCGC 180
DB 277 CATGTGGCGGACCTTTCTGCGAGGGAACCCCAACCTCAGTGTGAGATCTTTCACCGGCGC 336
QY 181 CTCTACTTCTGTGAGGACCGCAAGGCTGAGCCCGAGGCGCTGCGGCGCTGACCGCGCC 240
DB 337 CTCTACTTCTGTGAGGACCGCAAGGCTGAGCCCGAGGCGCTGCGGCGCTGACCGCGCC 396
QY 241 GGCGTGCAATAGCCATCATGACCTTCAAG 271
DB 397 GGCGTGCAATAGCCATCATGACCTTCAAG 427

RESULT 13
AF529835 596 bp mRNA linear SYN 17-SEP-2002
LOCUS Mus musculus clone 8 transgenic Homo sapiens AID (AID) mRNA,
DEFINITION complete cds.
ACCESSION AF529835
VERSION AF529835.1 GI:22297257
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE Martin, A. and Scharff, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
AUTHORS 2 (bases 1 to 596)
MARTIN, A. and SCHARFF, M.D.
TITLE Direct Submision
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA
FEATURES
source
1..596
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="8"
/cell_type="hybridoma PI-5"
/transgenic
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..>596
/gene="AID"
1..>596
/gene="AID"
/codon_start=1
/product="AID"
/protein_id="AA095422.1"
/db_xref="GI:22297258"
/translation="MDSLMNRKFLYQKRVNMAKGRRETYVYVYKRRDSATSFSL
DFGLRNKNGCHELLFLRYISMDLDPGRCYAVTWFTMSPCYDCAHVADFLRGNP
NLSIRITFALVYFCEDRKAPEGRLRHAGVOIALMTFKDYFCWNTFVENHRTFK
AMEGLHNSVRLSRQLRILLPLYEVDLDRARFTLGL"

ORIGIN
Query Match 100.0%; Score 271; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 5.9e-50;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCTGCACGCGAATTGCTTCTCGCTACATCTCGGACTGGACCTAGACCT 60
DB 157 AAGCGCTGCACGCGAATTGCTTCTCGCTACATCTCGGACTGGACCTAGACCT 216
QY 61 GGCGCTGTACCGCGTCACTGATTCACCTCTGTGAGCCCTGTGACGACTGTGCCCCGA 120

Db 217 GGGCGCTGCTACCGCGCTCACTGTTCACTCTGAGACCCCTGCTACACGACTGTGCGCGA 276

QY 121 CATGTGGCCGACTTTTCTGCGAGGAAACCCCACTGAGTGTGAGATCTTCAACCGCGCG 180

Db 277 CATGTGGCCGACTTTTCTGCGAGGAAACCCCACTGAGTGTGAGATCTTCAACCGCGCG 336

QY 181 CTCTACTTCTGTGAGACCGCGAGGCTGAGCGGGCTGCGGGCTGTGACACCGCGCC 240

Db 337 CTCTACTTCTGTGAGACCGCGAGGCTGAGCGGGCTGCGGGCTGTGACACCGCGCC 396

QY 241 GGGGTGCAATAGCATCATGACCTTCAAG 271

Db 397 GGGGTGCAATAGCATCATGACCTTCAAG 427

RESULT 14

AF529836 596 bp mRNA linear SYN 17-SEP-2002

LOCUS AF529836

DEFINITION Mus musculus clone 9 transgenic Homo sapiens AID (AID) mRNA,

complete cds.

ACCESSION AF529836

VERSION AF529836.1 GI:22297259

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 596)

AUTHORS Martin,A. and Scharff,M.D.

TITLE Somatic hypermutation of the AID transgene in B and non-B cells

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)

AUTHORS Martin,A. and Scharff,M.D.

TITLE Direct Submission

JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Ave, Chanin 404, Bronx, NY 10461, USA

FEATURES

source

1..596

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="9"

/cell_type="hybridoma PL-5"

/transgenic

1..596

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

1..596

/gene="AID"

1..596

/gene="AID"

/codon_start=1

/product="AID"

/protein_id="AA05423.1"

/db_xref="GI:22297260"

/translation="MDSLMNRKFLYQKRVNRKGRRETYLCYVTKRDSATPSFLDPGYLNKNGCHEVLEFLRYISDWMDLPGRCYRVWFTWSPCYCAHVAHDFLGNPNSLRIFTRALYFCEDRKAPBGLRLHRAQVQIAIMTKDYFCWNTFVNHERTFKAMEGLHENSVRLSRQLRIILPLYEVDLDRDAFRTGL"

ORIGIN

Query Match 100.0%; Score 271; DB 12; Length 596;

Best Local Similarity 100.0%; Pred. No. 5,9e-50;

Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGCGCTGCAAGTGAATTCCTCCCGCTACATCTCGGACTGGGACCTAGACCT 60

Db 157 AAGCGCTGCAAGTGAATTCCTCCCGCTACATCTCGGACTGGGACCTAGACCT 216

QY 61 GGGCGCTGCTACCGCGCTCACTGTTCACTCTGAGACCCCTGCTACACGACTGTGCGCGA 120

Db 217 GGGCGCTGCTACCGCGCTCACTGTTCACTCTGAGACCCCTGCTACACGACTGTGCGCGA 276

QY 121 CATGTGGCCGACTTTTCTGCGAGGAAACCCCACTGAGTGTGAGATCTTCAACCGCGCG 180

Db 277 CATGTGGCCGACTTTTCTGCGAGGAAACCCCACTGAGTGTGAGATCTTCAACCGCGCG 336

QY 181 CTCTACTTCTGTGAGACCGCGAGGCTGAGCGGGCTGCGGGCTGTGACACCGCGCC 240

Db 337 CTCTACTTCTGTGAGACCGCGAGGCTGAGCGGGCTGCGGGCTGTGACACCGCGCC 396

QY 241 GGGGTGCAATAGCATCATGACCTTCAAG 271

Db 397 GGGGTGCAATAGCATCATGACCTTCAAG 427

RESULT 15

BT007402 597 bp mRNA linear PRI 13-MAY-2003

LOCUS BT007402

DEFINITION Homo sapiens activation-induced cytidine deaminase mRNA, complete cds.

ACCESSION BT007402

VERSION BT007402.1 GI:30583642

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 597)

AUTHORS Kalline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y., Phelan,M., and Farmer,A.

TITLE Cloning of human full-length CDSs in BD Creator(TM) System Donor vector

JOURNAL Unpublished

AUTHORS 2 (bases 1 to 597)

Kalline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y., Phelan,M., and Farmer,A.

TITLE Direct Submission

JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA

COMMENT This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD in-Fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame. Clone distribution: <http://bioinfo.clontech.com/orfclones>.

FEATURES

source

1..597

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="GH00009X1.0"

/clone_1db="BD Creator(TM) CDS library derived from MGC collection"

/lab_host="DH5alpha TI resistant"

/notes="Vector: pDNR-Dual"

1..597

/codon_start=1

/product="activation-induced cytidine deaminase"

/protein_id="AAP3606.1"

/db_xref="GI:30583643"

/translation="MDSLMNRKFLYQKRVNRKGRRETYLCYVTKRDSATPSFLDPGYLNKNGCHEVLEFLRYISDWMDLPGRCYRVWFTWSPCYCAHVAHDFLGNPNSLRIFTRALYFCEDRKAPBGLRLHRAQVQIAIMTKDYFCWNTFVNHERTFKAMEGLHENSVRLSRQLRIILPLYEVDLDRDAFRTGL"

ORIGIN

Query Match 100.0%; Score 271; DB 9; Length 597;

Best Local Similarity 100.0%; Pred. No. 5,9e-50;

Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| Qy | 1 | AAGGCTGCCACGTGGAATTGCTCTCTCTCCGCTACATCTCGGACTGGGACCTAGACCT | 60 |
| Db | 157 | AAGGCTGCCACGTGGAATTGCTCTCTCTCCGCTACATCTCGGACTGGGACCTAGACCT | 216 |
| Qy | 61 | GGCCGCTGCTACCGGCTCACTGTTCACTCTCGAGCCCTGCTACGACTGTGCCGA | 120 |
| Db | 217 | GGCCGCTGCTACCGGCTCACTGTTCACTCTCGAGCCCTGCTACGACTGTGCCGA | 276 |
| Qy | 121 | CATGTGGCCGACTTTCTGCGAGGGAACCCCAACTCATGTGAGATCTTCAACCGCGGC | 180 |
| Db | 277 | CATGTGGCCGACTTTCTGCGAGGGAACCCCAACTCATGTGAGATCTTCAACCGCGGC | 336 |
| Qy | 181 | CTCTACTTCTGTAGAGACCGCAAGGCTGAGCCCGAGGGCTGGGGCGCTGCACCGCGCC | 240 |
| Db | 337 | CTCTACTTCTGTAGAGACCGCAAGGCTGAGCCCGAGGGCTGGGGCGCTGCACCGCGCC | 396 |
| Qy | 241 | GGGTCGCAATAGCCATCATGACCTTCAAAG | 271 |
| Db | 397 | GGGTCGCAATAGCCATCATGACCTTCAAAG | 427 |

Search completed: March 12, 2004, 21:22:31
Job time : 854.181 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 12:26:39 ; Search time 108.517 Seconds
(without alignments)
10609.014 Million cell updates/sec

Title: US-09-966-880A-13

Sequence: 1 aacggctgcacgtggaatt.....agccatcatgaccttcaag 271

Scoring table: IDENTITY NTC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|---------------------|
| 1 | 271 | 100.0 | 271 | 3 AAC55317 | AAC55317 Human act |
| 2 | 271 | 100.0 | 597 | 8 ADB88952 | ADB88952 AID gene |
| 3 | 271 | 100.0 | 1543 | 7 ABX05468 | ABX05468 Human nov |
| 4 | 271 | 100.0 | 2791 | 6 AB873287 | AB873287 DNA encod |
| 5 | 271 | 100.0 | 2791 | 6 AB873288 | AB873288 DNA encod |
| 6 | 271 | 100.0 | 2818 | 3 AAC55312 | AAC55312 Human act |
| 7 | 271 | 100.0 | 6564 | 3 AAC55314 | AAC55314 Human act |
| 8 | 271 | 100.0 | 11204 | 3 AAC55339 | AAC55339 Human act |
| 9 | 271 | 100.0 | 11204 | 6 AB873286 | AB873286 DNA encod |
| 10 | 221.6 | 81.8 | 2440 | 3 AAC55307 | AAC55307 Mouse act |
| 11 | 87.4 | 32.3 | 201143 | 2 ABR63568 | ABR63568 Human DNA |
| 12 | 85.8 | 31.7 | 610 | 2 AA4V8231 | AA4V8231 Human RNA |
| 13 | 85.8 | 31.7 | 610 | 3 AA4A72058 | AA4A72058 CDNA enco |
| 14 | 85.8 | 31.7 | 819 | 4 AA4A1659 | AA4A1659 CDNA enco |
| 15 | 85.8 | 31.7 | 944 | 6 AB199876 | AB199876 Human sec |
| 16 | 85.8 | 31.7 | 950 | 4 AA4A1420 | AA4A1420 CDNA enco |
| 17 | 85.8 | 31.7 | 987 | 4 AA4F16264 | AA4F16264 Human pro |
| 18 | 85.8 | 31.7 | 987 | 4 AA4A1407 | AA4A1407 CDNA enco |
| 19 | 85.8 | 31.7 | 1120 | 4 AA4I59847 | AA4I59847 Human pol |
| 20 | 85.8 | 31.7 | 1143 | 4 AA4I58061 | AA4I58061 Human pol |
| 21 | 85.8 | 31.7 | 1638 | 6 AB061187 | AB061187 Unidentif |
| 22 | 85.8 | 31.7 | 6131 | 4 AA4S42187 | AA4S42187 Genomic s |
| 23 | 85.8 | 31.7 | 6131 | 4 AA4K69784 | AA4K69784 Human imm |

| | | | | | |
|----|------|------|--------|-------------|---------------------|
| 24 | 85 | 31.4 | 1534 | 2 AA220856 | AA220856 Polynucle |
| 25 | 85 | 31.4 | 1534 | 4 AA4S59293 | AA4S59293 Human CDN |
| 26 | 85 | 31.4 | 1534 | 6 AB490962 | AB490962 Human pol |
| 27 | 85 | 31.4 | 1717 | 9 ADD18987 | ADD18987 Human dis |
| 28 | 85 | 31.4 | 2151 | 6 AA4S6572 | AA4S6572 CDNA sequ |
| 29 | 83.4 | 30.8 | 1446 | 3 AA4I2409 | AA4I2409 CDNA enco |
| 30 | 82.6 | 30.5 | 1055 | 6 AAD24392 | AA24392 Human RNA |
| 31 | 78.4 | 29.0 | 122748 | 6 ABT10719 | ABT10719 Human bre |
| 32 | 78.4 | 28.9 | 371 | 8 ACH43025 | ACH43025 Human foe |
| 33 | 77.4 | 28.6 | 1526 | 9 ADD18895 | ADD18895 Human dis |
| 34 | 76.8 | 28.3 | 823 | 9 ADE26568 | ADE26568 Human CDN |
| 35 | 76.8 | 28.3 | 1348 | 6 ABN96785 | ABN96785 Gene #328 |
| 36 | 73.4 | 27.1 | 1567 | 4 AA4I93080 | AA4I93080 Human gen |
| 37 | 70.2 | 25.9 | 716 | 2 AA4I1654 | AA4I1654 Human CDN |
| 38 | 68.4 | 25.2 | 2338 | 4 AA4I17654 | AA4I17654 Human CDN |
| 39 | 68.4 | 25.2 | 2773 | 9 ADB63605 | ADB63605 Human CDN |
| 40 | 64.8 | 23.9 | 572 | 4 AA4I8764 | AA4I8764 Probe #86 |
| 41 | 64.8 | 23.9 | 572 | 4 AB463759 | AB463759 Human foe |
| 42 | 64.8 | 23.9 | 572 | 4 AA4I3877 | AA4I3877 Probe #12 |
| 43 | 64.8 | 23.9 | 572 | 4 AB430950 | AB430950 Probe #94 |
| 44 | 64.8 | 23.9 | 572 | 4 AA4K3995 | AA4K3995 Human bon |
| 45 | 64.8 | 23.9 | 572 | 4 AA4I2273 | AA4I2273 Human bra |

ALIGNMENTS

| | |
|---|---|
| RESULT 1 | |
| AAC55317 | |
| ID AAC55317 standard; DNA; 271 BP. | |
| AC AAC55317; | |
| DT 05-FEB-2001 (first entry) | |
| DE Human activation-induced cytidine deaminase exon 3 SEQ ID NO:13. | |
| XX | Activation-induced cytidine deaminase; AID; cytidine deaminase; |
| KM | immune related disease; allergy; allergic disease; anti-allergic; |
| KM | antianemic; antidiabetic; ophthalmological; anti-HIV; dermatological; |
| KM | gene therapy; B cell associated immune system disorder; food allergy; |
| KM | immunodeficiency disease; immunoglobulin A deficiency disease; asthma; |
| KM | IgA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis; |
| KM | drug allergy; allergic rhinitis; Rosen disease; discoid dermatitis; AIDS; |
| KM | ataxia telangiectasia; common variable immunodeficiency disorder; |
| KM | major histocompatibility class II deficiency disease; |
| KM | auto immunodeficiency syndrome; Igs subclass selection disorder; ds. |
| OS Homo sapiens. | |
| XX | |
| XX | WO200058480-A1. |
| XX | |
| XX | 05-OCT-2000. |
| XX | |
| XX | 28-MAR-2000; 2000WO-JP001918. |
| XX | |
| XX | 29-MAR-1999; 99JP-00087192. |
| XX | |
| XX | 24-JUN-1999; 99JP-00178999. |
| XX | |
| XX | 27-DEC-1999; 99JP-00371382. |
| XX | |
| XX | (NTSB) JAPAN TOBACCO INC. |
| XX | (HONJ) HONJO T. |
| XX | |
| XX | Honjo T, Muramatsu M; |
| XX | |
| XX | WPI; 2000-611715/58. |
| XX | |
| XX | Nucleic acid encoding activation induced cytidine deaminase, useful as a |
| XX | target for drug development for immune-related diseases including |
| XX | allergies. |
| XX | |
| XX | Claim 18; Page 151; 174pp; Japanese. |

CC The present invention describes an activation-induced cytidine deaminase
CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
CC cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic,
CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
CC and can be used in gene therapy. AID polymucleotides are useful in
CC methods for identifying drugs for the treatment of B cell associated
CC immune system disorders, immunodeficiency diseases and allergies, such as
CC immunoglobulin A (IgA) deficiency disease, IgA nephritis, gamma-
CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
CC drug allergy, allergic rhinitis, Rosen disease, DiGeorge disease, ataxia
CC telangiectasia, common variable immunodeficiency disorder, MHC (major
CC histocompatibility class II deficiency disease, AIDS (auto
CC immunodeficiency syndrome), elevated IgE disorder, and IgG subclass
CC selection disorder. The DNA sequences encoding AID may be used for gene
CC therapy and the antibodies to the AID protein may be used for diagnosis
CC and treatment of these disorders. The present sequence represents the
CC exon 3 genomic DNA sequence of human AID
CC

SO Sequence 271 BP; 47 A; 95 C; 76 G; 53 T; 0 U; 0 Other;

Query Match 100.0%; Score 271; DB 3; Length 271;

Best Local Similarity 100.0%; Pred.No. 3.3e-58; Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGCTGCGACGTGGAAATGCTTCTCCCTGACATCTGGAGCTGGAGACCTAGACCT 60
DB 1 AACGGCTGCGACGTGGAAATGCTTCTCCCTGACATCTGGAGCTGGAGACCTAGACCT 60
QY 61 GGGCGGTGTACCGGGTACCGGTGACCGTCTGAGAGCCCTGTACAGCTGAGCCCGA 120
DB 61 GGGCGGTGTACCGGGTACCGGTGACCGTCTGAGAGCCCTGTACAGCTGAGCCCGA 120
QY 121 CATGTGCGCGACTTCTGCGAGAGAACCCCACTCAGTGTGAGGATCTTCAACCGCGCGC 180
DB 121 CATGTGCGCGACTTCTGCGAGAGAACCCCACTCAGTGTGAGGATCTTCAACCGCGCGC 180
QY 181 CTCTACTTCTGTGAGAGACCGCAAGGCTGAGCCCGAGGGCTGCGCGCTGACCGCGCC 240
DB 181 CTCTACTTCTGTGAGAGACCGCAAGGCTGAGCCCGAGGGCTGCGCGCTGACCGCGCC 240
QY 241 GGGGCGCAATAGCCATCATGACCTTCAAG 271
DB 241 GGGGCGCAATAGCCATCATGACCTTCAAG 271

RESULT 2

ADB88952 standard; DNA; 597 BP.

AC ADB88952;

DT 04-DEC-2003 (first entry)

DE AID gene as substrate used in AID mutagenic method.

KM ds; cytosstatic; virucide; antidote; mutation;
KM activation-induced cytidine deaminase; cell phenotype;
KM monoclonal antibody; antigen; cross-reactivity; tumour;
KM germ warfare agent; toxin; V region.

XX Homo sapiens.

OS Homo sapiens.

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT /tag= d
FT /note= "mutation in Pl-5 hybridoma cell"

FT /tag= e
FT /note= "mutation in Pl-5 hybridoma cell"

FT /tag= f
FT /note= "mutation in CHO cell"

FT /tag= g
FT /note= "mutation in Ramos cell"

FT /tag= h
FT /note= "mutation in Pl-5 hybridoma cell"

FT /tag= i
FT /note= "mutation in Ramos cell"

FT /tag= j
FT /note= "mutation in Pl-5 hybridoma cell"

FT /tag= k
FT /note= "mutation in Pl-5 hybridoma cell"

FT /tag= l
FT /note= "mutation in Pl-5 hybridoma cell"

FT /tag= m
FT /note= "mutation in Ramos cell"

FT /tag= n
FT /note= "mutation in Ramos cell"

FT /tag= o
FT /note= "mutation in CHO cell"

FT /tag= p
FT /note= "mutation in Ramos cell"

FT /tag= q
FT /note= "mutation in Ramos cell"

FT /tag= r
FT /note= "mutation in Ramos cell"

FT /tag= s
FT /note= "mutation in Ramos cell"

FT /tag= t
FT /note= "mutation in Ramos cell"

FT /tag= u
FT /note= "mutation in Ramos cell"

FT /tag= v
FT /note= "mutation in Ramos cell"

FT /tag= w
FT /note= "mutation in Ramos cell"

FT /tag= x
FT /note= "mutation in Ramos cell"

FT /tag= y
FT /note= "mutation in Ramos cell"

FT /tag= z
FT /note= "mutation in Pl-5 hybridoma cell"

FT /tag= aa
FT /note= "mutation in CHO cell"

FT /tag= ab
FT /note= "mutation in CHO cell"

| | | |
|----|---------------------------------|---|
| FT | mutation | /note= "mutation in PI-5 hybridoma cell" |
| FT | mutation | /replace (410,T) |
| FT | mutation | /*tag= ac |
| FT | mutation | /note= "mutation in PI-5 hybridoma cell" |
| FT | mutation | /replace (418,T) |
| FT | mutation | /*tag= ad |
| FT | mutation | /note= "mutation in PI-5 hybridoma cell" |
| FT | mutation | /replace (420,T) |
| FT | mutation | /*tag= ae |
| FT | mutation | /note= "mutation in PI-5 hybridoma cell" |
| FT | mutation | /replace (421,A) |
| FT | mutation | /*tag= af |
| FT | mutation | /note= "mutation in CHO cell" |
| FT | mutation | /replace (430,G) |
| FT | mutation | /*tag= ag |
| FT | mutation | /note= "mutation in Ramos cell" |
| FT | mutation | /replace (440,A) |
| FT | mutation | /*tag= ah |
| FT | mutation | /note= "mutation in CHO cell" |
| FT | mutation | /replace (449,T) |
| FT | mutation | /*tag= ai |
| FT | mutation | /note= "mutation in PI-5 hybridoma cell" |
| FT | mutation | /replace (495,A) |
| FT | mutation | /*tag= aj |
| FT | mutation | /note= "mutation in CHO cell" |
| FT | mutation | /replace (517,C) |
| FT | mutation | /*tag= ak |
| FT | mutation | /note= "mutation in Ramos cell" |
| FT | mutation | /replace (526,T, A) |
| FT | mutation | /*tag= al |
| FT | mutation | /note= "mutation in PI-5 hybridoma cell, A mutation in Ramos cell" |
| FT | mutation | /replace (533,A) |
| FT | mutation | /*tag= am |
| FT | mutation | /note= "mutation in PI-5 hybridoma cell" |
| FT | mutation | /replace (538,A) |
| FT | mutation | /*tag= an |
| FT | mutation | /note= "mutation in CHO cell" |
| FT | mutation | /replace (549,A) |
| FT | mutation | /*tag= ao |
| FT | mutation | /note= "mutation in CHO cell" |
| FT | mutation | /replace (568,T) |
| FT | mutation | /*tag= ap |
| FT | mutation | /note= "mutation in Ramos cell" |
| FT | mutation | /replace (570,T) |
| FT | mutation | /*tag= aq |
| FT | mutation | /note= "mutation in CHO cell" |
| XX | | WO2003061363-A2. |
| XX | | 31-JUL-2003. |
| XX | | 15-JAN-2003; 2003WO-US001149. |
| XX | | 17-JAN-2002; 2002US-0350269P. |
| XX | | (YESH) UNIV YESHIVA EINSTEIN COLLEGE. |
| PI | Martin A, Scharff MD; | |
| DR | WPI; 2003-747983/70. | |
| XX | | |
| XX | | |
| PT | | Induction of mutations in gene expressed in eukaryotic cell useful e.g. |
| PT | | to produce antibodies with greater affinity or specificity for antigens |
| PT | | e.g. human pathogens by expressing an activation-induced cytidine |
| PT | | deaminase gene in cell. |
| XX | | |
| PS | Example 4; Fig 6; 73p; English. | |
| CC | | The invention relates to the induction of mutations in a gene expressed |
| CC | | in a eukaryotic cell, where the gene is operably linked to a promoter and |
| CC | | within 2 kb of promoter, by expressing a transgenic activation-induced |
| CC | | cytidine deaminase (AID) gene in the cell. The method is useful for |

| Query Match | 100.0% | Score 271, | DB 8, | Length 597; |
|---|-----------------|--------------------|-----------|-------------|
| Best Local Similarity | 100.0% | Prid. No. 3.ee-56; | | |
| Matches 271, | Conservative 0, | Mismatches 0, | Indels 0, | Gaps 0, |
| QY 1 AAGCGCTGCCACGGTGAATTGCTCTTCCTCCGCTACATCTGGACTGGGAACTTGACCT 60 | | | | |
| Db 157 AAGCGCTGCCACGGTGAATTGCTCTTCCTCCGCTACATCTGGACTGGGAACTTGACCT 216 | | | | |
| QY 61 GGCGCGTCTACCCCGTCACTGGTTCACCTCTCGAGACCCCTGTAAGCTGACCGCGA 120 | | | | |
| Db 217 GGCGCGTCTACCCCGTCACTGGTTCACCTCTCGAGACCCCTGTAAGCTGACCGCGA 276 | | | | |
| QY 121 CATGTGGCCGACCTTCTTCGCGAGGAAACCCCAACTCACTGAGAGATCTTACCGCGGC 180 | | | | |
| Db 277 CATGTGGCCGACCTTCTTCGCGAGGAAACCCCAACTCACTGAGAGATCTTACCGCGGC 336 | | | | |
| QY 181 CTCTACTTCTGTGAGAACCGCAAGCTGAGGCCCGAGGGGCTGGCGGCTCAACCGGCC 240 | | | | |
| Db 337 CTCTACTTCTGTGAGAACCGCAAGCTGAGGCCCGAGGGGCTGGCGGCTCAACCGGCC 396 | | | | |
| QY 241 GGGGTGCAGATATAGCATCATGACCTTCAAG 271 | | | | |
| Db 397 GGGGTGCAGATATAGCATCATGACCTTCAAG 427 | | | | |

RESULT 3
ABX05468
ID ABX05468 standard; cDNA, 1543 BP.
AC ABX05468;
DT 17-JAN-2003 (first entry)
DE Human novel polynucleotide #483.
XX Human; gene; ss; genetic disorder; gene mapping; medical imaging; cancer;
KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;
KW fungal infection; bacterial infection; autoimmune disease; diabetes;
atopic dermatitis.
XX Homo sapiens.
OS
XX WO200274961-A1.
EN
XX 26-SEP-2002.
PD
XX 14-MAR-2002; 2002WO-US005109.
PF
XX 15-MAR-2001; 2001US-00810173.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao Q, Ren F;
P1 Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
P2 Wehrman T, Wang J, Wang D, Drmanac RT;
XX WPI; 2003-040556/03.
PR

DR P-PSDB; ABU00390.
XX
XX New isolated polypeptides and polynucleotides, useful for preventing,
PT treating or ameliorating medical conditions, such as cancer,
PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
PT disorders, and infections.
XX
XX Claim 1; SEQ ID NO 483; 235bp; English.
XX
XX The invention relates to human polynucleotides and the polypeptides they
CC encode. The polynucleotides and polypeptides are useful in diagnostics,
CC forensics, gene mapping, medical imaging, identification of mutations,
CC responsible for genetic disorders or other traits, assessing biodiversity
CC and producing many other types of data and products dependent on DNA and
CC amino acid sequences. They are also useful for preventing, treating or
CC ameliorating medical conditions, such as cancer, neurodegenerative
CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
CC Sequences ABX04986-ABX0511 represent human polynucleotides of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification but is based on sequence information supplied
CC by the European Patent Office
XX
SQ Sequence 1543 BP; 428 A; 334 C; 352 G; 423 T; 0 U; 0 Other;
Query Match 100.0%; Score 271; DB 7; Length 1543;
Best Local Similarity 100.0%; Pred. No. 4.2e-58;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCTGCCACGTGGAATGCTCTCCCTGATATCTGGACCTGGAACCTTACCT 60
DB 234 AAGCGCTGCCACGTGGAATGCTCTCCCTGATATCTGGACCTGGAACCTTACCT 293
QY 61 GGGCGCTGACCGCGACCGGCTTCACTCTGAGGCCCGCTGACGACCTGAGCCGA 120
DB 294 GGGCGCTGACCGCGCTTCTGAGGCCCGCTGAGGCCCGCTGACGACCTGAGCCGA 353
QY 121 CATGTGGCCGACTTTCTGCGAGGGAACCCCACTGAGTCTGAGATCTTCAACCGCGCG 180
DB 354 CATGTGGCCGACTTTCTGCGAGGGAACCCCACTGAGTCTGAGATCTTCAACCGCGCG 413
QY 181 CTCTACTTCTGTGAGACCGCAAGGCTGAGGCCGAGGGGCTGGCGCTGACCGCGCC 240
DB 414 CTCTACTTCTGTGAGACCGCAAGGCTGAGGCCGAGGGGCTGGCGCTGACCGCGCC 473
QY 241 GGGGTGCAATATGACCATCATGACCTTCAAG 271
DB 474 GGGGTGCAATATGACCATCATGACCTTCAAG 504

RESULT 4
ABX73287
ID ABX73287 standard; DNA; 2791 BP.
XX
XX ABX73287;
XX
XX 04-DEC-2002 (first entry)
XX
XX DNA encoding human translocation del(12p) protein #2.
XX
XX Chromosome aberration; oncogenic fusion protein; cancer;
XX proliferative disease; cellular protein isoform; heat shock protein 90;
XX HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
XX T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
XX acute myeloid leukaemia; ALL; APL; NHL; solid tumour;
XX papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
XX rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.
XX
XX Homo sapiens.
XX

PN WO200269900-A2.
XX
XX 12-SEP-2002.
XX
XX 01-MAR-2002; 2002MO-US006518.
XX
XX 01-MAR-2001; 2001US-0272751P.
XX
XX (CONF-) CONFORMA THERAPEUTICS CORP.
XX
XX Fritz LC, Burrows FJ;
XX
XX WPI; 2002-698710/75.
DR P-PSDB; ABG95083.
XX
XX Treating genetically-defined disease associated with chromosomal
PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
PT diseases, involves administering an inhibitor of heat shock protein 90.
XX
XX Disclosure; Page 246-247; 389pp; English.
XX
XX The invention describes a method of treating genetically-defined disease
CC associated with chromosomal aberrations yielding oncogenic fusion
CC proteins (I), treating cancerous cells containing (I) in a heterogeneous
CC cell population, treating proliferative diseases associated with mutant
CC protein or cellular protein isoforms (II) dependent on heat shock protein
CC (HSP)-90, or selectively treating cells expressing (II) involving
CC administering HSP90-inhibitor. The method is useful for treating
CC genetically-defined disease with chromosomal aberration yielding
CC oncogenic fusion protein, treating cancerous cells containing fusion
CC protein in heterogeneous cell population, treating proliferative disease
CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
CC p53), or selectively treating cells expressing mutant protein or cellular
CC protein isoform in a patient heterozygous for (II). The method is useful
CC for treating a disease e.g. haematopoietic disorder such as T or B cell
CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CML,
CC or a disease characterised by a solid tumour such as papillary thyroid
CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
CC synovial sarcoma. The method is also useful for treating viral
CC infections. This represents the DNA sequence of a chromosome aberration
XX
SQ Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;
Query Match 100.0%; Score 271; DB 6; Length 2791;
Best Local Similarity 100.0%; Pred. No. 4.5e-58;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCTGCCACGTGGAATGCTCTCCCTGATATCTGGACCTGGAACCTTACCT 60
DB 233 AAGCGCTGCCACGTGGAATGCTCTCCCTGATATCTGGACCTGGAACCTTACCT 292
QY 61 GGGCGCTGACCGCGCTTCTGCGAGGGAACCCCACTGAGTCTGAGATCTTCAACCGCGCG 120
DB 293 GGGCGCTGACCGCGCTTCTGCGAGGGAACCCCACTGAGTCTGAGATCTTCAACCGCGCG 352
QY 121 CATGTGGCCGACTTTCTGCGAGGGAACCCCACTGAGTCTGAGATCTTCAACCGCGCG 180
DB 353 CATGTGGCCGACTTTCTGCGAGGGAACCCCACTGAGTCTGAGATCTTCAACCGCGCG 412
QY 181 CTCTACTTCTGTGAGACCGCAAGGCTGAGGCCGAGGGGCTGGCGCTGACCGCGCC 240
DB 413 CTCTACTTCTGTGAGACCGCAAGGCTGAGGCCGAGGGGCTGGCGCTGACCGCGCC 472
QY 241 GGGGTGCAATATGACCATCATGACCTTCAAG 271
DB 473 GGGGTGCAATATGACCATCATGACCTTCAAG 503

RESULT 5
ABX73288
ID ABX73288 standard; DNA; 2791 BP.
XX

AC ABS73288;
 XX
 DT 04-DEC-2002 (first entry)
 XX
 DE DNA encoding human translocation del(12p) protein #3.
 XX
 KW Chromosome aberration; oncogenic fusion protein; cancer;
 KM proliferative disease; cellular protein isoform; heat shock protein 90;
 KM HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
 KW T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
 KW acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
 KW acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
 KM papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
 KM rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200269900-A2.
 XX
 PD 12-SEP-2002.
 XX
 PF 01-MAR-2002; 2002WO-US006518.
 XX
 PR 01-MAR-2001; 2001US-0272751P.
 XX
 PA (CONF-) CONFORMA THERAPEUTICS CORP.
 XX
 FI Fritz LC, Burrows FJ;
 XX
 DR WPI: 2002-698710/75.
 DR P-PSDB; ABG95084.
 XX
 PT Treating genetically-defined disease associated with chromosomal
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
 PT diseases, involves administering an inhibitor of heat shock protein 90.
 XX
 PS Disclosure; Page 248-249; 389pp; English.
 XX
 CC The invention describes a method of treating genetically-defined disease
 CC associated with chromosomal aberrations yielding oncogenic fusion
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous
 CC cell population, treating proliferative diseases associated with mutant
 CC protein or cellular protein isoforms (II) dependent on heat shock protein
 CC (HSP)-90, or selectively treating cells expressing (II) involving
 CC administering HSP90-inhibitor. The method is useful for treating
 CC genetically-defined disease with chromosomal aberration yielding
 CC oncogenic fusion protein, treating cancerous cells containing fusion
 CC protein in heterogeneous cell population, treating proliferative disease
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
 CC p53), or selectively treating cells expressing mutant protein or cellular
 CC protein isoform in a patient heterozygous for (II). The method is useful
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, AML, NHL and CMML,
 CC or a disease characterised by a solid tumour such as papillary thyroid
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
 CC synovial sarcoma. The method is also useful for treating viral
 CC infections. This represents the DNA sequence of a chromosome aberration
 XX
 SO Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;
 Query Match 100.0%; Score 271; DB 6; Length 2791;
 Best Local Similarity 100.0%; Pred. No. 4.5e-58;
 Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AACGGCTGACACGCGAATGCTTCTTCCTCGCTACACTCTCGAGCTGGAGCCTTAGACCT 60
 DB 233 AACGGCTGACACGCGAATGCTTCTTCCTCGCTACACTCTCGAGCTGGAGCCTTAGACCT 292
 QY 61 GGGCGCTGACACGCGAATGCTTCTTCCTCGCTACACTCTCGAGCTGGAGCCTTAGACCT 120
 DB 293 GGGCGCTGACACGCGAATGCTTCTTCCTCGCTACACTCTCGAGCTGGAGCCTTAGACCT 352

QY 121 CATGTGGCGAATTCTTCTGCGAGGAGAACCCCACTCAGTCTGAGAACTTTCACCGCGCC 180
 DB 353 CATGTGGCGAATTCTTCTGCGAGGAGAACCCCACTCAGTCTGAGAACTTTCACCGCGCC 412
 QY 181 CTCTACTTCTGTGAGAGACCGCAAGGCTGAGGGGCTGCGGCGGCTGACCGGGCC 240
 DB 413 CTCTACTTCTGTGAGAGACCGCAAGGCTGAGGGGCTGCGGCGGCTGACCGGGCC 472
 QY 241 GGGGTGCAAAATGACATCATGACCTTCAAG 271
 DB 473 GGGGTGCAAAATGACATCATGACCTTCAAG 503
 RESULT 6
 AAC55312
 ID AAC55312 standard; cDNA; 2818 BP.
 XX
 AC AAC55312;
 XX
 DT 05-FEB-2001 (first entry)
 XX
 DE Human activation-induced cytidine deaminase encoding cDNA SEQ ID NO:7.
 XX
 KW Activation-induced cytidine deaminase; AID; cytidine deaminase;
 KW immune related disease; allergy; allergic disease; anti-allergic;
 KW antianemic; antiasthmatic; ophthalmological; anti-HIV; dermatological;
 KW gene therapy; B cell associated immune system disorder; food allergy;
 KW immunodeficiency disease; immunoglobulin A deficiency disease; aschma;
 KW IGA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;
 KW drug allergy; allergic rhinitis; Rosen disease; Digeorge disease; AIDS;
 KW ataxia telangiectasia; common variable immunodeficiency disorder;
 KW major histocompatibility class II deficiency disease;
 KW auto immunodeficiency syndrome; IgG subclass selection disorder; ss.
 KW
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 80..676
 FT /*tag= a
 FT /product= "activation-induced cytidine deaminase"
 PN WO200058480-A1.
 PD 05-OCT-2000.
 PF 28-MAR-2000; 2000WO-JP001918.
 PR 29-MAR-1999; 99JP-00087192.
 PR 24-JUN-1999; 99JP-00178999.
 PR 27-DEC-1999; 99JP-00371382.
 PA (NIBS) JAPAN TOBACCO INC.
 PA (HONJ/) HONJO T.
 PI Honjo T, Muramatsu M;
 DR WPI: 2000-611715/58.
 DR P-PSDB; AAB24198.
 PT Nucleic acid encoding activation induced cytidine deaminase, useful as a
 PT target for drug development for immune-related diseases including
 PT allergies.
 PS Claim 3; Page 135-139; 174pp; Japanese.
 XX
 CC The present sequence encodes human activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic,
 CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-

CC globulinaemia, atopic dermatitis, allergic colitis, asthma, food allergy,
CC drug allergy, allergic rhinitis, Rosen disease, DiGeorge disease, ataxia
CC telangiectasia, common variable immunodeficiency disorder, MHC (major
CC histocompatibility class II deficiency disease, AIDS (auto
CC immunodeficiency syndrome), elevated IgE disorder, and IgG subclass
CC selection disorder. The DNA sequences encoding AID may be used for gene
CC therapy and the antibodies to the AID protein may be used for diagnosis
CC and treatment of these disorders
XX
SQ Sequence 2818 BP; 868 A; 548 C; 626 G; 776 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 271; DB 3; Length 2818;
Best Local Similarity 100.0%; Pred. No. 4,5e-58;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACGGCTGCCAGTGGAAATGCTCTTCCTCCGTACATCTGGAGCTGGAGCTAGACCTT 60
DB 236 AACGGCTGCCAGTGGAAATGCTCTTCCTCCGTACATCTGGAGCTGGAGCTAGACCTT 295
QY 61 GGGCGCTGTACCGCGCTACCTGATTCACCTCTGGAGCCCTGTACACTGTGGCCGA 120
DB 296 GGGCGCTGTACCGCGCTACCTGATTCACCTCTGGAGCCCTGTACACTGTGGCCGA 355
QY 121 CATGTGGCCGACCTTTCTGCGAGGGAACCCCAACCTAGTCTGAGATCTTACCCGGCGC 180
DB 356 CATGTGGCCGACCTTTCTGCGAGGGAACCCCAACCTAGTCTGAGATCTTACCCGGCGC 415
QY 181 CTCTACTCTGTGAGGAGCGGAGGCTGAGCCGAGGCGCTCGGGGCTGAGACCGCGCC 240
DB 416 CTCTACTCTGTGAGGAGCGGAGGCTGAGCCGAGGCGCTCGGGGCTGAGACCGCGCC 475
QY 241 GGGGTGCAATATGACCATCATGACCTTCAAG 271
DB 476 GGGGTGCAATATGACCATCATGACCTTCAAG 506
RESULT 7
AAC55314
ID AAC55314 standard; DNA; 6564 BP.
AC AAC55314;
XX
DT 05-FEB-2001 (first entry)
XX
DE Human activation-induced cytidine deaminase genomic DNA SEQ ID NO:10.
XX
KW Activation-induced cytidine deaminase; AID; cytidine deaminase;
KW immune related disease; allergy; allergic disease; antiallergic;
KW antianaemic; antiaesthatic; ophthalmological; anti-HIV; dermatological;
KW gene therapy; B cell associated immune system disorder; food allergy;
KW immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
KW IgA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis;
KW drug allergy; allergic rhinitis; Rosen disease; DiGeorge disease; AIDS;
KW ataxia telangiectasia; common variable immunodeficiency disorder;
KW major histocompatibility class II deficiency disease;
KW auto immunodeficiency syndrome; IgG subclass selection disorder; ds.
XX
OS Homo sapiens.
XX
XX WO200058480-A1.
XX
PD 05-OCT-2000.
XX
PF 28-MAR-2000; 2000WO-JP001918.
XX
XX 29-MAR-1999; 99JP-00087192.
PR 24-JUN-1999; 99JP-00178999.
PR 27-DEC-1999; 99JP-00371382.
XX
XX (NIBS) JAPAN TOBACCO INC.
PA (HONW/) HONJO T.
XX
PI Honjo T, Muramatsu M;

XX
DR WPI; 2000-611715/58.
XX
XX Nucleic acid encoding activation induced cytidine deaminase, useful as a
PT target for drug development for immune-related diseases including
PT allergies.
XX
PS Claim 17; Page 145-150; 174pp; Japanese.
XX
XX The present invention describes an activation-induced cytidine deaminase
XX (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
XX cytidine activity similar to APOBEC-1. AID has antiallergic, antianaemic,
XX antiaesthatic, ophthalmological, anti-HIV and dermatological activities,
XX and can be used in gene therapy. AID polynucleotides are useful in
XX methods for identifying drugs for the treatment of B cell associated
XX immune system disorders, immunodeficiency diseases and allergies, such as
XX immunoglobulin A (IgA) deficiency disease, IgA nephritis, gamma-
XX globulinaemia, atopic dermatitis, allergic colitis, asthma, food allergy,
XX drug allergy, allergic rhinitis, Rosen disease, DiGeorge disease, ataxia
XX telangiectasia, common variable immunodeficiency disorder, MHC (major
XX histocompatibility class II deficiency disease, AIDS (auto
XX immunodeficiency syndrome), elevated IgE disorder, and IgG subclass
XX selection disorder. The DNA sequences encoding AID may be used for gene
XX therapy and the antibodies to the AID protein may be used for diagnosis
XX and treatment of these disorders. The present sequence represents a
XX genomic DNA sequence of human AID
SQ Sequence 6564 BP; 1909 A; 1358 C; 1383 G; 1914 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 271; DB 3; Length 6564;
Best Local Similarity 100.0%; Pred. No. 5.1e-58;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACGGCTGCCAGTGGAAATGCTCTTCCTCCGTACATCTGGAGCTGGAGCTAGACCTT 60
DB 2592 AACGGCTGCCAGTGGAAATGCTCTTCCTCCGTACATCTGGAGCTGGAGCTAGACCTT 2651
QY 61 GGGCGCTGTACCGCGCTACCTGATTCACCTCTGGAGCCCTGTACACTGTGGCCGA 120
DB 2652 GGGCGCTGTACCGCGCTACCTGATTCACCTCTGGAGCCCTGTACACTGTGGCCGA 2711
QY 121 CATGTGGCCGACCTTTCTGCGAGGGAACCCCAACCTAGTCTGAGATCTTACCCGGCGC 180
DB 2712 CATGTGGCCGACCTTTCTGCGAGGGAACCCCAACCTAGTCTGAGATCTTACCCGGCGC 2771
QY 181 CTCTACTCTGTGAGGAGCGGAGGCTGAGCCGAGGCGCTCGGGGCTGAGACCGCGCC 240
DB 2772 CTCTACTCTGTGAGGAGCGGAGGCTGAGCCGAGGCGCTCGGGGCTGAGACCGCGCC 2831
QY 241 GGGGTGCAATATGACCATCATGACCTTCAAG 271
DB 2832 GGGGTGCAATATGACCATCATGACCTTCAAG 2862
RESULT 8
AAC55339
ID AAC55339 standard; DNA; 11204 BP.
AC AAC55339;
XX
DT 05-FEB-2001 (first entry)
XX
DE Human activation-induced cytidine deaminase genomic DNA SEQ ID NO:35.
XX
KW Activation-induced cytidine deaminase; AID; cytidine deaminase;
KW immune related disease; allergy; allergic disease; antiallergic;
KW antianaemic; antiaesthatic; ophthalmological; anti-HIV; dermatological;
KW gene therapy; B cell associated immune system disorder; food allergy;
KW immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
KW IgA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis;
KW drug allergy; allergic rhinitis; Rosen disease; DiGeorge disease; AIDS;
KW ataxia telangiectasia; common variable immunodeficiency disorder;
KW major histocompatibility class II deficiency disease;

KM auto immunodeficiency syndrome; IgG subclass selection disorder; ds.
XX Homo sapiens.
XX
XX MO200058480-A1.
XX
XX PD 05-OCT-2000.
XX
XX PF 28-MAR-2000; 2000MO-JP001918.
XX
XX PR 29-MAR-1999; 99DP-00087192.
XX PR 24-JUN-1999; 99DP-00178999.
XX PR 27-DEC-1999; 99DP-00371382.
XX
XX (NHSB) JAPAN TOBACCO INC.
XX (HONN/) HONTO T.
XX
XX Honjo T, Muramatsu M;
XX
XX WPI; 2000-611715/58.
XX
XX PT Nucleic acid encoding activation induced cytidine deaminase, useful as a
XX target for drug development for immune-related diseases including
XX allergies.
XX
XX PS Claim 17; Page 163-170; 174pp; Japanese.
XX
XX CC The present invention describes an activation-induced cytidine deaminase
XX (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
XX cytidine activity similar to APOBEC-1. AID has anti-allergic, anti-nausea,
XX antispasmodic, ophthalmological, anti-HIV and dermatological activities,
XX and can be used in gene therapy. AID polynucleotides are useful in
XX methods for identifying drugs for the treatment of B cell associated
XX immune system disorders, immunodeficiency diseases and allergies, such as
XX immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
XX globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
XX drug allergy, allergic rhinitis, allergic disease, Digorge disease, ataxia
XX telangiectasia, common variable immunodeficiency disorder, MHC (major
XX histocompatibility class II deficiency disease, AIDS (auto
XX immunodeficiency syndrome), elevated IGE disorder, and IgG subclass
XX selection disorder. The DNA sequences encoding AID may be used for gene
XX therapy and the antibodies to the AID protein may be used for diagnosis
XX and treatment of these disorders. The present sequence represents a
XX genomic DNA sequence of human AID
XX
XX SQ Sequence 11204 BP; 3305 A; 2273 C; 2373 G; 3253 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 271; DB 3; Length 11204;
XX Best Local Similarity 100.0%; Pred. No. 5,5e-58;
XX Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 AACGGCTGACAGTGAATGCTCTCTCCGCTACATCTCGAGTGGAGCTAGACCT 60
XX DB 7807 AACGGCTGACAGTGAATGCTCTCTCCGCTACATCTCGAGTGGAGCTAGACCT 7866
XX
XX QY 61 GGGCGGTGCTACCGCGCTACCTGTTACCTCTGAGGCCCTGCTACGACTGTGCCGA 120
XX DB 7867 GGGCGGTGCTACCGCGCTACCTGTTACCTCTGAGGCCCTGCTACGACTGTGCCGA 7926
XX
XX QY 121 CATGTGGCGCACTTTCTGAGAGGAAACCCCACTTACGTGAGATCTTCAACGGCGGC 180
XX DB 7927 CATGTGGCGCACTTTCTGAGAGGAAACCCCACTTACGTGAGATCTTCAACGGCGGC 7986
XX
XX QY 181 CTGTACTTGTGAGAGCGGCAAGCTAGGCCCGGAGGGGCTGGGCGGTGAGACGGCGGC 240
XX DB 7987 CTGTACTTGTGAGAGCGGCAAGCTAGGCCCGGAGGGGCTGGGCGGTGAGACGGCGGC 8046
XX
XX QY 241 GGGGTCAATATAGCATCATGACCTTCAAG 271
XX DB 8047 GGGGTCAATATAGCATCATGACCTTCAAG 8077
XX
XX RESULT 9

AB573286
XX ID AB573286 standard; DNA; 11204 BP.
XX
XX AC AB573286;
XX
XX DT 04-DEC-2002 (first entry)
XX
XX DE DNA encoding human translocation del(12p) protein #1.
XX
XX
XX KM Chromosome aberration; oncogenic fusion protein; cancer;
XX proliferative disease; cellular protein isoform; heat shock protein 90;
XX HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
XX T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
XX acute myeloid leukaemia; ALL; chronic myelomonocytic leukaemia; CMML;
XX acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
XX papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
XX rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.
XX
XX OS Homo sapiens.
XX
XX PM WO200269900-A2.
XX
XX PD 12-SEP-2002.
XX
XX PF 01-MAR-2002; 2002MO-US006518.
XX
XX PR 01-MAR-2001; 2001US-0272751P.
XX
XX (CONF-) CONFORMA THERAPEUTICS CORP.
XX
XX PA Fritzt LC, Burrows FU;
XX
XX PI WPI; 2002-698710/75.
XX
XX DR P-PSDB; ABG95082.
XX
XX PT Treating genetically-defined disease associated with chromosomal
XX aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
XX diseases, involves administering an inhibitor of heat shock protein 90.
XX
XX PS Disclosure; Page 242-245; 389pp; English.
XX
XX CC The invention describes a method of treating genetically-defined disease
XX associated with chromosomal aberrations yielding oncogenic fusion
XX proteins (I), treating cancerous cells containing (I) in a heterogeneous
XX cell population, treating proliferative diseases associated with mutant
XX protein or cellular protein isoforms (II) dependent on heat shock protein
XX (HSP)-90, or selectively treating cells expressing (II) involving
XX administering HSP90-inhibitor. The method is useful for treating
XX genetically-defined disease with chromosomal aberration yielding
XX oncogenic fusion protein, treating cancerous cells containing fusion
XX protein in heterogeneous cell population, treating proliferative disease
XX (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
XX cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
XX p53), or selectively treating cells expressing mutant protein or cellular
XX protein isoform in a patient heterozygous for (II). The method is useful
XX for treating a disease e.g. haematopoietic disorder such as T or B cell
XX lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CMML,
XX or a disease characterised by a solid tumour such as papillary thyroid
XX carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
XX synovial sarcoma. The method is also useful for treating viral
XX infections. This represents the DNA sequence of a chromosome aberration
XX
XX SQ Sequence 11204 BP; 3305 A; 2273 C; 2373 G; 3253 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 271; DB 6; Length 11204;
XX Best Local Similarity 100.0%; Pred. No. 5,5e-58;
XX Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 AACGGCTGACAGTGAATGCTCTCTCCGCTACATCTCGAGTGGAGCTAGACCT 60
XX DB 7807 AACGGCTGACAGTGAATGCTCTCTCCGCTACATCTCGAGTGGAGCTAGACCT 7866
XX
XX QY 61 GGGCGGTGCTACCGCGCTACCTGTTACCTCTGAGGCCCTGCTACGACTGTGCCGA 120

```

Db      7867 GGGCCCTGCTACCGGCTGACCTGTTCACTCTGAGACCCCTGTACGACTGTGCGCGA 7926
QY      121 CATGTGGCCGCACTTTCTGTGGAGGAAACCCCACTCAGTGTGAGATTTTACCGCGCGC 180
Db      7927 CATGTGGCCGCACTTTCTGTGGAGGAAACCCCACTCAGTGTGAGATTTTACCGCGCGC 7986
QY      181 CTCTACTCTGTGAGGACCGCAAGCTGAGCCCGAGGCGCTGCGCGCTGACCGCGCC 240
Db      7987 CTCTACTCTGTGAGGACCGCAAGCTGAGCCCGAGGCGCTGCGCGCTGACCGCGCC 8046
QY      241 GGGGTGCATAATAGCCATCATGACCTTCAAG 271
Db      8047 GGGGTGCATAATAGCCATCATGACCTTCAAG 8077

RESULT 10
AAC5307
ID      AAC5307 standard; cDNA; 2440 BP.
XX      AAC5307;
AC      AAC5307;
XX      05-FEB-2001 (first entry)
DE      Mouse activation-induced cytidine deaminase encoding cDNA SEQ ID NO:1.
XX      Activation-induced cytidine deaminase; AID; cytidine deaminase;
XX      immune related disease; allergy; allergic disease; antiallergic;
XX      antianemic; antisthmatic; ophthalmological; anti-HIV; dermatological;
XX      gene therapy; B cell associated immune system disorder; food allergy;
XX      immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
XX      IgA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;
XX      drug allergy; allergic rhinitis; Rosen disease; Disgeorge disease; AIDS;
XX      ataxia telangiectasia; common variable immunodeficiency disorder;
XX      major histocompatibility class II deficiency disease;
XX      auto immunodeficiency syndrome; Ig sublcass selection disorder; ss.
XX      Mus musculus.
XX      Mus musculus.
XX      Key      Location/Qualifiers
XX      CDS      93..689
XX      FT      /*tag= a
XX      FT      /product= "activation-induced cytidine deaminase"
XX      FT      /product= "activation-induced cytidine deaminase"
XX      PN      WO200058480-A1.
XX      PD      05-OCT-2000.
XX      PF      28-MAR-2000; 2000WO-0P001918.
XX      PR      29-MAR-1999; 99JP-00087192.
XX      PR      24-JUN-1999; 99JP-00178999.
XX      PR      27-DEC-1999; 99JP-00371382.
XX      PA      (NISR ) JAPAN TOBACCO INC.
XX      PA      (HONJ/) HONJO T.
XX      PI      Honjo T, Muramatsu M;
XX      DR      WPI; 2000-611715/58.
XX      DR      P-PSDB; AAB24197.
XX      PT      Nucleic acid encoding activation induced cytidine deaminase, useful as a
XX      PT      target for drug development for immune-related diseases including
XX      PT      allergies.
XX      PS      Claim 3; Page 126-130; 174pp; Japanese.
XX      CC      The present sequence encodes mouse activation-induced cytidine deaminase
XX      CC      (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
XX      CC      cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic,
XX      CC      antisthmatic, ophthalmological, anti-HIV and dermatological activities,
XX      CC      and can be used in gene therapy. AID polynucleotides are useful in

```

```

CC      methods for identifying drugs for the treatment of B cell associated
CC      immune system disorders, immunodeficiency diseases and allergies, such as
CC      immunoglobulin A (IGA) deficiency disease, IgA nephritis, gamma-
CC      globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
CC      drug allergy, allergic rhinitis, Rosen disease, Disgeorge disease, ataxia
CC      telangiectasia, common variable immunodeficiency disorder, MHC (major
CC      histocompatibility class) class II deficiency disease, AIDS (auto
CC      immunodeficiency syndrome), elevated IgE disorder, and IgG subclass
CC      selection disorder. The DNA sequences encoding AID may be used for gene
CC      therapy and the antibodies to the AID protein may be used for diagnosis
CC      and treatment of these disorders
XX      Sequence 2440 BP; 706 A; 546 C; 551 G; 636 T; 0 U; 1 Other;
XX      Query Match      81.8%; Score 221.6; DB 3; Length 2440;
XX      Best Local Similarity 89.2%; Pred. No. 9.2e-46;
XX      Matches 239; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY      4 GGGTGCACGTGAGATTTGCTCTCTCCCTACATCTCGAGTGGAGACCTAGACCTTGC 63
Db      252 GGGTGCACGTGAGATTTGCTCTCTCCCTACATCTCGAGTGGAGACCTAGACCTTGC 311
QY      64 CGCTGTACCGGCTGACCTGTGCTTACCTCTGAGAGCCCTGCTAGACGTGCGGACAT 123
Db      312 CGGTGTACCGGCTGACCTGTGCTTACCTCTGAGAGCCCTGCTAGACGTGCGGACAT 371
QY      124 GTGGCCGACCTTTCTGTGGAGGAAACCCCACTCAGTGTGAGATTTTACCGCGCGCTC 183
Db      372 GTGGCCGACCTTTCTGTGGAGGAAACCCCACTCAGTGTGAGATTTTACCGCGCGCTC 431
QY      184 TACTTCTGTGAGAGCGCAAGCGTGAAGCGGAGGCTCGGCGGTGACCGCGCGCTG 243
Db      432 TACTTCTGTGAGAGCGCAAGCGTGAAGCGGAGGCTCGGCGGTGACCGCGCGCTG 491
QY      244 GTGCAATATAGCCATCATGACCTTCAAG 271
Db      492 GTGCAATATAGCCATCATGACCTTCAAG 519

RESULT 11
ABK83568
ID      ABK83568 standard; DNA; 201143 BP.
XX      ABK83568;
AC      ABK83568;
XX      29-AUG-2002 (first entry)
DE      Human DNA differentially expressed in granulocytic cells #139.
XX      Human DNA differentially expressed in granulocytic cells #139.
XX      Human; de; granulocytic cell; DNA chip; bacterial infection;
XX      viral infection; parasitic infection; protozoal infection;
XX      fungal infection; sterile inflammatory disease; psoriasis;
XX      rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
XX      cardiac reperfusion injury; renal reperfusion injury; ARDS;
XX      adult respiratory distress syndrome; inflammatory bowel disease;
XX      Crohn's disease; ulcerative colitis; periodontal disease;
XX      granulocyte activation; chronic inflammation; allergy.
XX      Homo sapiens.
XX      WO200228999-A2.
XX      11-APR-2002.
XX      PD      03-OCT-2001; 2001WO-US030821.
XX      PF      03-OCT-2000; 2000US-0237189P.
XX      PR      03-OCT-2000; 2000US-0237189P.
XX      PA      (GENE-) GENE LOGIC INC.
XX      PA      Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;
XX      WPI; 2002-435328/46.

```

XX Detecting granulocyte activation by detecting differential expression of
PT genes associated with granulocyte activation, which serves as diagnostic
PT markers that is useful for monitoring disease states and drug toxicity.

XX Claim 1; SEQ ID NO 139; 114bp; English.

XX The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comprising the
CC expression level to an expression level in an unactivated GC, where
CC differential expression of Gs is indicative of GCA. Also included are
CC modulating (M2) GCA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences

XX Sequence 201143 BP; 50606 A; 49308 C; 49683 G; 51546 T; 0 U; 0 Other;

XX Query Match 32.3%; Score 87.4; DB 6; Length 201143;

XX Best Local Similarity 60.8%; Pred. No. 4.9e-12;

XX Matches 161; Conservative 0; Mismatches 101; Indels 3; Gaps 1;

XX 7 TGCCAGGTGGAATTGCTCTTCCCTCCGCTACATCTCGAGCTGGAGCACTTACATCTTCAACACA 167958
DB 167899 TGTCAATGCAAGAAAGTCTCTCTCTGTTCTGTGAGACATCTTCTCTTCAACACA 167958

XX 67 TGCTACCGGCTCACTGCTTCACTCTCTGAGGCGCTGTACGACATGTGCCCCGACATGTG 126

XX DB 167959 AACTACGAGGTCACTGCTGATCAATCTTGGAGCCCTTCCGAGAGTGTGAGGAGGTG 168018

XX 127 GCCGACTTCTGCGAGGAAACCCCAACTCAGTCTGAGATCTTACCCGCGGCTCTAC 186

XX DB 168019 GCCGAGTCTCTGCGAGGACACCAAGTAATCTCAACATCTTACCCGCGGCTCTAC 168078

XX 187 TTCTGTGAGGACCGCAAGGTGAGCCGAGGGCTGCGGCGGTGACACCGCGCGGGGTG 246

XX DB 168079 TACTTCTGGA---TCAAGATTACCGAGAGGAGGCTCCGAGCTGAGTGAAGAAAGGGGCTC 168135

XX 247 CAATAGCCATGACCTTCAAG 271

XX DB 168136 TCCGTGAGATCATGGCTACAAAG 168160

XX RESULT 12
XX AAV48231
XX ID AAV48231 standard; cDNA; 610 BP.

XX AAV48231;

XX 09-NOV-1998 (first entry)

XX Human RNA editing enzyme nucleotide sequence.

XX ss: human; RNA editing enzyme; REE; pharmaceutical carrier; cancer;
XX viral diseases; circulatory system disorder; RNA processing; B.
XX hypercholesterolemia; alpha-galactosidase; apolipoprotein B.
XX Homo sapiens.

XX Key Location/Qualifiers
XX CDS 1..573
XX /*tag= a
XX /product= "RNA editing enzyme"

XX US5804185-A.

XX 08-SEP-1998.

XX 13-MAR-1997; 97US-00816241.

XX 13-MAR-1997; 97US-00816241.

XX (INCY-) INCYTE PHARM INC.

XX Bandman O, Gold SK,

XX WPI; 1998-505585/43.

XX P-PSDB; AAW77092.

XX Human RNA editing enzyme and poly(nucleotide(s) encoding it - useful for
XX recombinant production of the enzyme and treatment and detection of
XX disorders associated with incorrect RNA processing.

XX Claim 4; Fig 1; 27pp; English.

XX The human RNA editing enzyme (REE) is used in a pharmaceutical carrier
XX for the treatment of cancer, viral diseases and circulatory system
XX disorders. The enzyme is used in vivo for the correct processing of RNA
XX transcripts of genes e.g. change of a codon in apolipoprotein B (apoB)
XX RNA to give a 100 and 48 kDa product transcribed from the same gene.
XX Certain disorders have been linked to incorrect RNA editing, e.g. failure
XX of apo B editing leads to excessive apoB 100 production and
XX hypercholesterolemia. Other disorders thought to be linked to incorrect
XX RNA processing include aberrant alpha-galactosidase processing in Fabry's
XX disease and neurofibromatosis type I. The enzyme can be produced
XX recombinantly to treat related disorders. It can also be used to raise
XX antibodies for immuno-based detection of REE expression levels e.g. ELISA

XX Sequence 610 BP; 158 A; 154 C; 154 G; 144 T; 0 U; 0 Other;

XX Query Match 31.7%; Score 85.8; DB 2; Length 610;

XX Best Local Similarity 60.4%; Pred. No. 5.3e-12;

XX Matches 160; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

XX 7 TGCCAGGTGGAATTGCTCTTCCCTCCGCTACATCTCGAGCTGGAGCACTTACCTTGGCGGC 66

XX DB 193 TGTCAATGCAAGAAAGTCTCTCTCTGTTCTGTGAGACACATCTTCTCTTCAACACA 252

XX 67 TGCTACCGGCTCACTGCTTCACTCTCTGAGGCGCTGTACGACATGTGCCCCGACATGTG 126

XX DB 253 AAGTACCAAGGTCACTGCTGATCAATCTTGGAGCCCTTCCGAGAGTGTGAGGAGGTG 312

XX 127 GCCGACTTCTGCGAGGAAACCCCAACTCAGTCTGAGATCTTACCCGCGGCTCTAC 186

XX DB 313 GCCGAGTCTCTGCGAGGACACCAAGTAATCTCAACATCTTACCCGCGGCTCTAC 372

XX 187 TTCTGTGAGGACCGCAAGGTGAGCCGAGGGCTGCGGCGGTGACACCGCGCGGGGTG 246

XX DB 373 TACTTCCAGTATC---CATGTTCACAGGAGGGGCTCCGAGCTGAGTGAAGAAAGGGGTC 429

QY 247 CAATAGCCATGACCTTCAAAG 271
 Db 430 GCTGTGAGATCATGACTATGAG 454

RESULT 13
 AAA72058
 ID AAA72058 standard; cDNA; 610 BP.
 XX AAA72058;

XX 24-NOV-2000 (first entry)

XX cDNA encoding human RNA editing enzyme REE-2.

XX RNA editing enzyme; REE-2; human; HEPR homologue; REPR homologue;
 KM photoboln I homologue; cancer; tumour; autoimmune disorder;
 KM circulatory system disorder; hypercholesterolaemia; viral infection;
 KM neurological disease; neurofibromatosis; transcript editing; detection;
 KM ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..573

XX /tag= a
 XX /product= "Human REE-2"

XX US6087108-A.

XX 11-JUL-2000.

XX 03-AUG-1998; 98US-00128395.

XX 13-MAR-1997; 97US-00816241.

XX (INCY-) INCYTE PHARM INC.

XX Goli SK, Bandman O;

XX WPI; 2000-531340/48.

XX P-PSDB; AAB11973.

XX Detecting polynucleotide encoding human RNA editing enzyme comprising
 PT hybridizing an isolated and purified polynucleotide complementary to the
 PT polynucleotide and detecting the hybridization complex.

XX Claim 5; Fig 1A-B; 27pp; English.

XX This sequence represents the cDNA encoding human RNA editing enzyme REE-
 CC 2. cDNA encoding REE-2 was initially isolated in a prostate tumour cDNA
 CC library, with the present sequence representing a consensus. REE-2 has
 CC chemical and structural homology with the human apob mRNA editing protein
 CC HEPR (28% identity), the rat HEPR homologue REPR (30% identity), and a
 CC portion of the mRNA editing enzyme photoboln I (43% identity). REE-2 was
 CC found to be expressed in a variety of cDNA libraries, a high proportion
 CC of which were derived from tumours, neuronal tissues, immune system cells
 CC or synovial tissue from arthritis patients. REE-2 is therefore thought to
 CC be associated with the development of cancer, autoimmune disorders,
 CC circulatory system disorders (e.g. hypercholesterolaemia), viral
 CC infections and neurological diseases (e.g., neurofibromatosis). REE-2 or
 CC its nucleic acids may be used in the diagnosis, treatment and prevention
 CC of such diseases via the modulation of transcript editing, which in turn
 CC has effects on the encoded protein (e.g., an alteration in protein
 CC activity). The invention specifically relates to methods of detecting
 CC nucleic acids encoding human REE-2 in a biological sample

XX Sequence 610 BP; 158 A; 154 C; 154 G; 144 T; 0 U; 0 Other;

XX Query Match 31.7%; Score 85.8; DB 3; Length 610;
 XX Best Local Similarity 60.4%; Pred. No. 5,3e-12;
 XX Matches 160; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

QY 7 TGGACGTGGAATTTGCTCTTCCCGCTACATGAGACTGGAGACCTAGACCTGGACCG 66
 Db 193 TGTATATGAGAAAGGTGCTCTCTCTTGGTTTGAGACACTACTGTCTCTTAACA 252
 QY 67 TGGTACCGCGTCACTGTTCACTCTGGAGCCCTGTACAGACTGTGCCGACATGTG 126
 Db 253 AAGTACAGGTCACTGTGACATCTTGGAGCCCTTGGCCAGACTGTGAGGGAGGTG 312
 QY 127 GCCGACTTTTGGGAGGAAACCCAACTTCAGTGAAGATTTTCAACCGCGGCTTAC 186
 Db 313 GCCGAGTTCCTGGCCAGGACAGCAAGTGAATCTCACCATTCTTCAACCGCGCTTAC 372
 QY 187 TTCTGTGAGAGACCGCAAGGCTGAGCCCGAGGGGCTGGCGGTGACCCGCGGGGTG 246
 Db 373 TACTTCCAGTATC--CAGTTTACAGAGAGGGGCTCCGAGCCTGAGTACAGAGAGGGGTG 429
 QY 247 CAATAGCCATGACCTTCAAAG 271
 Db 430 GCTGTGAGATCATGACTATGAG 454

RESULT 14
 AAS41669
 ID AAS41669 standard; cDNA; 819 BP.
 XX AAS41669;

XX 17-DEC-2001 (first entry)

XX cDNA encoding novel human enzyme polypeptide #895.

XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
 KM ligase; hyperproliferative disorder; immunodeficiency disorder;
 KM autoimmune disorder; neurological disorder; metabolic disorder;
 KM inflammatory disorder; cardiovascular disorder; reproductive disorder;
 KM blood-related disorder; infectious disorder; gene therapy; cytostatic;
 KM anti arthritic; nephroretropic; anticoagulant; ss.

XX Homo sapiens.

XX WO200155301-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001239.

XX 31-JAN-2000; 2000US-0179065P.
 XX 04-FEB-2000; 2000US-0180628P.
 XX 24-FEB-2000; 2000US-0184664P.
 XX 02-MAR-2000; 2000US-0186550P.
 XX 16-MAR-2000; 2000US-0189874P.
 XX 17-MAR-2000; 2000US-0190076P.
 XX 18-APR-2000; 2000US-0198123P.
 XX 19-MAY-2000; 2000US-020515P.
 XX 07-JUN-2000; 2000US-0209467P.
 XX 28-JUN-2000; 2000US-0214866P.
 XX 30-JUN-2000; 2000US-021513P.
 XX 07-JUL-2000; 2000US-0216647P.
 XX 07-JUL-2000; 2000US-0216880P.
 XX 11-JUL-2000; 2000US-0217487P.
 XX 11-JUL-2000; 2000US-0217496P.
 XX 14-JUL-2000; 2000US-0218290P.
 XX 26-JUL-2000; 2000US-0220963P.
 XX 26-JUL-2000; 2000US-0220964P.
 XX 14-AUG-2000; 2000US-0224518P.
 XX 14-AUG-2000; 2000US-0224519P.
 XX 14-AUG-2000; 2000US-0225213P.
 XX 14-AUG-2000; 2000US-0225214P.
 XX 14-AUG-2000; 2000US-0225266P.
 XX 14-AUG-2000; 2000US-0225267P.
 XX 14-AUG-2000; 2000US-0225268P.
 XX 14-AUG-2000; 2000US-0225270P.

Best Local Similarity 60.4%; Pred. No. 5.6e-12;
Matches 160; Conservative 0; Mismatches 102; Indels 3; Gaps 1

| | | | |
|----|-----|---|-----|
| Qy | 67 | TGCTACCGCGCTCACTGCTTCACTCTGGAAGCCCTGTCACGACTGTGCCCCGACATGTG | 126 |
| Db | 291 | TGTCTATGCAAGAAAGTGCTTCCCTCTCTTGTGTTCTTGCGACGACATACACTGTCTCTTCAACA | 350 |
| Qy | 127 | TGCGACGTGGAATTTGCTTCTCTCCCTCCGCTACATCTCGAAGTGGAACTGAACTTGCGCGC | 66 |
| Db | 351 | AGGTACCAAGTCTCACTGGTATACATCTTTGAAAGCCCTTGCCAGACTGTGCAAGGAAAGTG | 410 |
| Qy | 411 | GCGACCTTTTGTGAGGAAAGCCCAACCTCACTGTGAAGATTTCAACCGCGAGCTGTAC | 186 |
| Db | 470 | GCGAAGTTCCTGGCCAGGACGACGAAACGTGAATTTCAACATTTTACCGCCCGGCTTAC | 470 |
| Qy | 187 | TTCTGTGAGGACCGCAAGGCTGAGCCCGAGGAGCTGTGCGAGGCTGACACCGCGCCGGAGTG | 246 |
| Db | 471 | TACTTCCAGTATC---CATGTTACCGAGAGGGGCTCCGACAGCTGAGTCAGAAAGGGGCTC | 527 |
| Qy | 247 | CAAAATGCGCATGATGACCTTCAAG | 271 |
| Db | 528 | GCTGTGAGATCATGACCTATGAAAG | 552 |

RESULT 15

ID ABL99876 standard; cDNA; 944 BP.

AC ABL99876;

DT 03-OCT-2002 (first entry)

Human secretory polynucleotide (sptm) 131

KM Human; gene; secretory protein; secretory polynucleotides; SPM;
 KM SPM-related disease; somatic gene therapy; germ-line gene therapy;
 KM severe combined immunodeficiency; intracellular parasite protection;
 KM fungal parasite; protozoan parasite; cell proliferative disorder; cancer;
 KM immune disorder; AIDS; neurological disorder; Parkinson's disease;
 KM motor neuron disorder; demyelinating disease; multiple sclerosis;
 KM meningitis; abscess; prion disease; cerebral palsy;
 KM neuroskeletal disorder; peripheral nervous system disorder;
 KM dermatomyositis; polymyositis; myopathy; myasthenia gravis;
 KM mental disorder; Jourette's syndrome.

OS Homo sapiens.

PN WO200220756-A2.

PD 14-MAR-2002.

30-AUG-2001; 2001MO-US027297.

PR 05-SEP-2000; 2000US-0229747P.
PR 05-SEP-2000; 2000US-0229748P.
PR 05-SEP-2000; 2000US-0229749P.
PR 05-SEP-2000; 2000US-0229750P.
PR 05-SEP-2000; 2000US-0229751P.
PR 05-SEP-2000; 2000US-0230016P.
PR 05-SEP-2000; 2000US-0230583P.
PR 05-SEP-2000; 2000US-0230505P.
PR 06-SEP-2000; 2000US-0230514P.
PR 06-SEP-2000; 2000US-0230515P.
PR 06-SEP-2000; 2000US-0230517P.
PR 06-SEP-2000; 2000US-0230518P.
PR 06-SEP-2000; 2000US-0230519P.
PR 06-SEP-2000; 2000US-0230595P.
PR 06-SEP-2000; 2000US-0230596P.
PR 06-SEP-2000; 2000US-0230597P.
PR 06-SEP-2000; 2000US-0230599P.
PR 06-SEP-2000; 2000US-0230610P.
PR 06-SEP-2000; 2000US-0230664P.
PR 06-SEP-2000; 2000US-0230865P.

PR 06-SEP-2000; 2000US-0230988P.
PR 06-SEP-2000; 2000US-0230988P.
PR 06-SEP-2000; 2000US-0230992P.
PR 07-SEP-2000; 2000US-0230896P.
PR 07-SEP-2000; 2000US-0230897P.
PR 07-SEP-2000; 2000US-0230951P.
PR 07-SEP-2000; 2000US-023163P.
PR 07-SEP-2000; 2000US-0231833P.
XX (INCY-) INCYTE GENOMICS INC.
PA

PA (INCY-) INCYTE GENOMICS INC.

PI Stuart J., Lincoln SE, Altus CM, Dufour G, Chalup MS, Hillman JL;
PI Jones AL, Yu CY, Wright RJ, Gersen D, Lu TF, Yap PE, Dail CR;
PI Momiyama MG, Bradley DL, Robarici SD, Harris B, Roseberry AM;
PI Gerstman EH, Peralta CH, David MH, Panzer SR, Flores V, Daffio A,
PI Marwana R, Chen AJ, Chiang SC, Au AP, Imman KR;

DR WPI; 2002-315658/35.
DR P-PSDB; ABB97879.

PT. Polynucleotide sequences encoding human secretory proteins useful for gene therapy of e.g. genetic deficiency disorders, cancers, and diseases caused by intracellular parasites.

PS Claim 1; Page 326; 585pp; English

CC The invention comprises the amino acid and coding sequences of human
CC secretory (SPM) proteins. The SPM DNA and amino acid sequences are
CC useful for treating a disease or condition associated with the expression
CC of functional SPM. The SPM DNA sequences are useful for somatic or
CC germline gene therapy to correct a genetic deficiency (e.g. severe
CC combined immunodeficiency). The SPM DNA sequences are also useful in
CC providing protection against intracellular parasites (e.g. fungal
CC parasites and protozoan parasites). The SPM DNA and protein sequences
CC are also useful for diagnosing cell proliferative disorders, cancer,
CC immune disorders (e.g. AIDS), neurological disorders (e.g. Parkinson's
CC disease), motor neuron disorders, demyelinating diseases (e.g. multiple
CC sclerosis), meningitis, abscesses, prion diseases, cerebellar palsy,
CC neuroblastoma disorders, peripheral nervous system disorders,
CC dermatomyositis and polymyositis, myopathy, myasthenia gravis, and mental
CC disorders (e.g. Tourette's syndrome). cDNA sequences ABL97746 - ABL99929
CC represent human secretory polynucleotides of the invention
CX
XK
XQ Sequence 944 BP, 222 A, 284 C, 219 G, 219 T, 0 U, 0 Other;

Query Match 31.7%; Score 85.8; DB 6; Length 944;

Matches 160; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

[illegible]

Search completed: March 12, 2004, 18:08:05
Job time : 110.517 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 17:39:25 ; Search time 20.1637 Seconds

(without alignments)
7458.540 Million cell updates/sec

Title: US-09-966-880A-13

Perfect score: 271

Sequence: 1 aacggctgcacacgtggaatt.....agccatcatgacttcaag 271

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 27747546 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PCUTS.COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 85.8 | 31.7 | 610 | 1 | US-08-816-241-2 |
| 2 | 85.8 | 31.7 | 610 | 3 | US-09-128-395-2 |
| 3 | 50.6 | 18.7 | 879 | 1 | US-08-158-682A-3 |
| 4 | 44 | 16.2 | 891 | 1 | US-08-687-895-2 |
| 5 | 44 | 16.2 | 891 | 2 | US-09-040-483-2 |
| 6 | 41 | 15.1 | 879 | 1 | US-08-158-682A-1 |
| 7 | 41 | 15.1 | 879 | 1 | US-08-015-203-1 |
| 8 | 38 | 14.0 | 1929 | 4 | US-09-380-420C-1 |
| 9 | 38 | 14.0 | 1929 | 4 | US-09-899-642A-1 |
| 10 | 37.8 | 13.9 | 922 | 4 | US-09-189-060B-73 |
| 11 | 36.4 | 13.4 | 624 | 4 | US-09-252-991A-10053 |
| 12 | 36.4 | 13.4 | 1959 | 4 | US-09-252-991A-9783 |
| 13 | 35.8 | 13.2 | 4689 | 3 | US-09-105-537-34 |
| 14 | 35.8 | 13.2 | 36778 | 3 | US-09-105-537-5 |
| 15 | 35.8 | 13.2 | 38506 | 3 | US-09-320-878-19 |
| 16 | 35.8 | 13.2 | 38506 | 3 | US-09-141-908-1 |
| 17 | 35.8 | 13.2 | 38506 | 4 | US-09-657-440-19 |
| 18 | 35.4 | 13.1 | 2190 | 4 | US-09-015-188-1 |
| 19 | 35.2 | 13.0 | 2584 | 4 | US-09-716-129-47 |
| 20 | 34.6 | 12.8 | 1098 | 4 | US-09-170-496D-117 |
| 21 | 34.6 | 12.8 | 1098 | 4 | US-09-170-496D-225 |
| 22 | 34.6 | 12.8 | 1597 | 2 | US-08-724-974A-1 |
| 23 | 34.6 | 12.8 | 1697 | 4 | US-09-364-425B-26 |
| 24 | 34.4 | 12.7 | 2575 | 4 | US-09-716-129-25 |
| 25 | 34.4 | 12.7 | 2828 | 4 | US-09-016-434-148 |
| 26 | 34.4 | 12.7 | 2980 | 4 | US-09-266-225D-11 |
| 27 | 34 | 12.5 | 2505 | 4 | US-09-334-818A-1 |

| | | | | | | |
|------|------|------|---------|---|----------------------|-------------------|
| C 28 | 33.8 | 12.5 | 417 | 4 | US-09-252-991A-9297 | Sequence 9297, Ap |
| C 29 | 33.8 | 12.5 | 1215 | 4 | US-09-252-991A-9259 | Sequence 9259, Ap |
| C 30 | 33.8 | 12.5 | 1488 | 4 | US-09-252-991A-9307 | Sequence 9307, Ap |
| C 31 | 33.8 | 12.5 | 1602 | 4 | US-09-252-991A-9270 | Sequence 9270, Ap |
| C 32 | 33.8 | 12.5 | 15872 | 3 | US-09-105-537-1 | Sequence 1, Appli |
| C 33 | 33.8 | 12.5 | 15872 | 4 | US-09-091-609-1 | Sequence 1, Appli |
| C 34 | 33.8 | 12.5 | 15872 | 4 | US-09-091-609-3 | Sequence 3, Appli |
| C 35 | 33.8 | 12.5 | 4403765 | 3 | US-09-103-840A-2 | Sequence 2, Appli |
| C 36 | 33.8 | 12.5 | 4411529 | 3 | US-09-103-840A-1 | Sequence 1, Appli |
| C 37 | 33.6 | 12.4 | 1683 | 4 | US-09-252-991A-8337 | Sequence 8337, Ap |
| C 38 | 33.6 | 12.4 | 1833 | 4 | US-09-252-991A-8371 | Sequence 8371, Ap |
| C 39 | 33.6 | 12.4 | 1980 | 4 | US-09-252-991A-8507 | Sequence 8507, Ap |
| C 40 | 33 | 12.2 | 4911 | 4 | US-09-718-852-1 | Sequence 1, Appli |
| C 41 | 33 | 12.2 | 4911 | 4 | US-09-718-852-1 | Sequence 1, Appli |
| C 42 | 33 | 12.2 | 4911 | 4 | US-09-718-852-1 | Sequence 1, Appli |
| C 43 | 32.8 | 12.1 | 1017 | 4 | US-09-252-991A-10831 | Sequence 10831, A |
| C 44 | 32.6 | 12.0 | 1356 | 1 | US-08-317-880-4 | Sequence 4, Appli |
| C 45 | 32.6 | 12.0 | 1356 | 2 | US-08-782-396-4 | Sequence 4, Appli |

ALIGNMENTS

RESULT 1
US-08-816-241-2
Sequence 2, Application US/08816241
Patent No. 5804185
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
TITLE OF INVENTION: NOVEL RNA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/816,241
FILING DATE: Filed Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0239 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSTUT09
CLONE: 1646833
US-08-816-241-2
Query Match 31.7%; Score 85.8; DB 1; Length 610;
Best Local Similarity 60.4%; Pred. No. 1,6e-102;
Matches 160; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

QY 7 TGCACGCGAATGCTCTTCCTCCGCTACATCTCGAGTGGAGCCTAGACCTCGGCGC 66
DB 193 TGTATGAGAAAGGTGCTTCTCTCTGTTCTGCGACACATCTGTCTCTTAACA 252
QY 67 TGTATGCGGCTGACCTGTTGCTCACTCTGAGAGCCCTGCTACAGACTGTGCCGACATGTG 126
DB 253 AAGTACCGAGGTACCTGTGTACACATCTTGAGAGCCCTTGCCAGACTGTGAGGGAGGTG 312
QY 127 GCCGACTTCTGCGAGGGAACCCCACTCACTGAGAGATCTTCAACCGCGGCTCTTAC 186
DB 313 GCCGAGTCTCTGCGAGGGAACCCCACTCACTGAGAGATCTTCAACCGCGGCTCTTAC 312
QY 187 TTTGTGAGAGACCGGAGGCTGAGCGCGAGGGGCTGCGAGCGGCGCGGGGTG 246
DB 373 TACTTCCAGTATC--CATGTACAGAGAGGGGCTCCGACGCTGAGTCAAGAAAGGGGTC 429
QY 247 CAATAGCCATCATGACCTTCAAG 271
DB 430 GCTGTGAGATCATGACTATGAG 454

RESULT 2

US-09-128-395-2
Sequence 2, Application US/09128395
Patent No. 6087108
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL RNA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/128.395
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/816,241
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION/DOCKET NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0239 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSTITUT09
CLONE: 1646833
US-09-128-395-2

Query Match 31.7%; Score 85.8; DB 3; Length 610;
Best Local Similarity 60.4%; Pred. No. 1,6e-12;
Matches 160; Conservative 0; Mismatches 102; Indels 3; Gaps 1;
QY 7 TCCCACTGGAATGCTCTTCTCCGCTACATCTCGAGTGGAGCCTAGACCTTGGCGC 66

DB 193 TGTATGAGAAAGGTGCTTCTCTCTGTTCTGCGACACATCTGTCTCTTAACA 252
QY 67 TGTATGCGGCTGACCTGTTGCTCACTCTGAGAGCCCTGCTACAGACTGTGCCGACATGTG 126
DB 253 AAGTACCGAGGTACCTGTGTACACATCTTGAGAGCCCTTGCCAGACTGTGAGGGAGGTG 312
QY 127 GCCGACTTCTGCGAGGGAACCCCACTCACTGAGAGATCTTCAACCGCGGCTCTTAC 186
DB 313 GCCGAGTCTCTGCGAGGGAACCCCACTCACTGAGAGATCTTCAACCGCGGCTCTTAC 372
QY 187 TTTGTGAGAGACCGGAGGCTGAGCGCGAGGGGCTGCGAGCGGCGCGGGGTG 246
DB 373 TACTTCCAGTATC--CATGTACAGAGAGGGGCTCCGACGCTGAGTCAAGAAAGGGGTC 429
QY 247 CAATAGCCATCATGACCTTCAAG 271
DB 430 GCTGTGAGATCATGACTATGAG 454

RESULT 3

US-08-158-682A-3
Sequence 3, Application US/08158682A
Patent No. 5434058
GENERAL INFORMATION:
APPLICANT: Davidson, Nicholas O.
TITLE OF INVENTION: Apolipoprotein B RNA Editing Protein:
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: 321 No. 5434058th Clark Street, Suite 800
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,682A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Cooley, Ronald B.
REGISTRATION/DOCKET NUMBER: 27,187
REFERENCE/DOCKET NUMBER: ARCD:085
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 744-0090
TELEFAX: (312) 245-4961
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 879 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 23..730
US-08-158-682A-3

Query Match 18.7%; Score 50.6; DB 1; Length 879;
Best Local Similarity 59.3%; Pred. No. 0.00065;
Matches 86; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 40 TCCGACTGGAGACCTTAGACCTTGCGGCTGCTACCGGCTGACCTGTGCTACCTCTGAGAC 99
DB 236 TCGAAGAGATTTTACCCATCATCTGAGAGCTGACATCACTGCTGTCTGAGAGT 295
QY 100 CCTGTACGACTGTGCGGAGCATGTGGCGGACTTTCTGCGAGGGAACCCCACTCACT 159

? COUNTRY: USA
 ? ZIP: 60610
 ?
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: Floppy disk
 ? COMPUTER: IBM PC compatible
 ? OPERATING SYSTEM: PC-DOS/MS-DOS
 ? SOFTWARE: PatentIn Release #1.0, Version #1.25
 ?
 ? CURRENT APPLICATION DATA:
 ? APPLICATION NUMBER: US/08/158,682A
 ? FILING DATE:
 ?
 ? CLASSIFICATION: 435
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: Coolley, Ronald B.
 ? REGISTRATION NUMBER: 27,187
 ? REFERENCE/DOCKET NUMBER: ARCD:085
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: (312) 744-0090
 ? TELEFAX: (312) 245-4961
 ? INFORMATION FOR SEQ ID NO: 1:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 879 base pairs
 ? TYPE: nucleic acid
 ? STRANDEDNESS: single
 ? TOPOLOGY: linear
 ? MOLECULE TYPE: DNA (genomic)
 ? US-08-158-682A-1

| | | |
|--------------------------|--|---------------------------------------|
| | REGISTRATION NUMBER: | 27,187 |
| | REFERENCE/DOCKET NUMBER: | ARCD-069 |
| | TELECOMMUNICATION INFORMATION: | |
| | TELEPHONE: | (312) 744-0090 |
| | TELEFAX: | (312) 245-4961 |
| | INFORMATION FOR SEQ ID NO: | 1: |
| | SEQUENCE CHARACTERISTICS: | |
| | LENGTH: | 879 base pairs |
| | TYPE: | nucleic acid |
| | STRANDEDNESS: | single |
| | TOPOLOGY: | linear |
| | MOLECULE TYPE: | DNA (genomic) |
| | US-08-015-203-1 | |
| Query Match | 15.1%; | Score 41; DB 1; Length 879; |
| Best Local Similarity | 55.2%; | Pred. No. 0.15; 65; Indels 0; Gaps 0; |
| Matches 80; Conservative | 0; Mismatches | 0; Indels 0; Gaps 0; |
| QY | 41 CGGACTGGGACCTAGACCCTGAGCGCGTCAACGCCTCACCTGGTTCACTCCTCTGAGGCC 100 | |
| DB | 255 CAGAAAGATACTTTTGTCCTAACACCCAGATGCTCTCATTAAGTGTGCTCTGTGAAGTC 314 | |
| QY | 101 CCTGCTACGACGTGTGCCGACATGTGGCGCATTTCTTGCGAGGGAAACCCCAAAGCTCAGTC 160 | |
| DB | 315 CCTGTGGGAGTGCTCCAGGGCCATTACAGATTTTTGAAGCGATAACCCCATGTAAATC 374 | |
| QY | 161 TGAGATCTTCACGCGCGCCTCTA 185 | |
| DB | 375 TGTTTATTATATAGCAGCGGCTTTA 399 | |

FEATURE:
NAME/KEY: CDS
LOCATION: 81..1673
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-380-420C-1

Query Match 14.0%; Score 38; DB 4; Length 1929;
Best Local Similarity 53.3%; Pred. No. 0.86;
Matches 80; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 80 CCTGTTCACTCTCTGAGACCCCTGCTACGACTGTGCCGACATGTGCGCACTTTCTGC 139
DB 793 CCAGCTTCTCCGCGGAGACTTCTTCCCAACGCCGCGCGGCGCTCGCCGACCGCTCT 852
QY 140 GAGGAACCCCACTCACTGAGATTTTCAACCGCGGCGCTTACTTTGTAGAGACC 199
DB 853 CGGGCTTCTCGCGCGCGCGGACGATCTTCAACGAGCTGACGCTTTCTTGAGAGG 912
QY 200 GCAAGGCTGAGCCCGAGGGGCTCGCGGCGC 229
DB 913 TCAATGACCAAGCATGACCGCGCGCGCC 942

RESULT 9

US-09-899-642A-1
Sequence 1, Application US/09899642A
Patent No. 6649814
GENERAL INFORMATION:
APPLICANT: Halkier, Barbara
Bak, Soren

Kahn, Rachel
Moller, Birger
TITLE OF INVENTION: Cytochrome P450 Monooxygenases
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSER: Syngenta Patent Dept.
STREET: 3054 Cornwallis Road
CITY: RTP
STATE: NC
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,642A
FILING DATE: 05-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/380,420
FILING DATE: 12-No. 6649814-1999
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-21251A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1929 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: P450ox
FEATURE:
NAME/KEY: CDS
LOCATION: 81..1673
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-899-642A-1

Query Match 14.0%; Score 38; DB 4; Length 1929;
Best Local Similarity 53.3%; Pred. No. 0.86;
Matches 80; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 80 CCTGTTCACTCTCTGAGACCCCTGCTACGACTGTGCCGACATGTGCGCACTTTCTGC 139
DB 793 CCAGCTTCTCCGCGGAGACTTCTTCCCAACGCCGCGCGGCGCTCGCCGACCGCTCT 852
QY 140 GAGGAACCCCACTCACTGAGATTTTCAACCGCGGCGCTTACTTTGTAGAGACC 199
DB 853 CGGGCTTCTCGCGCGCGGACGATCTTCAACGAGCTGACGCTTTCTTGAGAGG 912
QY 200 GCAAGGCTGAGCCCGAGGGGCTCGCGGCGC 229
DB 913 TCAATGACCAAGCATGACCGCGCGCGCC 942

RESULT 10

US-09-189-060B-73
Sequence 73, Application US/09189060B
Patent No. 6270968
GENERAL INFORMATION:
APPLICANT: Dalboge, Henrik
Sandal, Thomas
APPLICANT: Kauppinen, Markus
Borge, Diderichsen

TITLE OF INVENTION: Method Of Providing No. 6270968e1 DNA Sequences
FILE REFERENCE: 4772.204-US
CURRENT APPLICATION NUMBER: US/09/189,060B
CURRENT FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: PCT/DK97/00216
PRIOR FILING DATE: 1997-05-12

NUMBER OF SEQ ID NOS: 74
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 73
LENGTH: 922
TYPE: DNA
ORGANISM: Hybrid

FEATURE:
NAME/KEY: CDS
LOCATION: (2)...(922)
US-09-189-060B-73

Query Match 13.9%; Score 37.8; DB 3; Length 922;
Best Local Similarity 49.3%; Pred. No. 0.9;
Matches 99; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 20 TGCCTTCTCCGCTACATCTCGGACTGGAGCTAGACCTGAGCCGCTGTACCGGTCA 79
DB 378 TGAATGCTCAGTCCACCAACGCGGCGGACTGTCCGACCAACCTTGAGCTCTCA 437
QY 80 CCGGTTCACTCTCTGAGACCCCTGCTACGACTGTGCCGACATGTGCGCACTTTCTGC 139
DB 438 TCCCGGCGCGCGGCGGAGATTTTCAACCGCGGCGCTTACTTTGTAGAGACC 497
QY 140 GAGGAACCCCACTCACTGAGATTTTCAACCGCGGCGCTTACTTTGTAGAGACC 199
DB 498 CGGGCGGCGGAGTACGAGGCGGTGAGCTCCGCGCGGAGTGCACGATGCCGAGGCCA 557
QY 200 GCAAGGCTGAGCCCGAGGGGCTCGCGGCGC 220
DB 558 TCAAGCGCGGCTGCGAGTGGC 578

RESULT 11

US-09-252-991A-10053/c
Sequence 10053, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO: 10053
LENGTH: 624
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10053

Query Match 13.4%; Score 36.4; DB 4; Length 624;
Best Local Similarity 51.9%; Pred. No. 1.9;
Matches 82; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 87 CACCTCCTGAGCCCTGCTACAGCTGTGCGGACATGTGGCCGACTTTCTGCGAGGAA 146
DB 367 CAACCTCACTCTCTGAGCTGACCTTACCCGCACTGAGGTGCGCTTGACCGGAT 308
QY 147 CCCCACTCAGTGTGAGGATTTCAACGCGCGCTTACTTCTGTGAGAACCGCAAGC 206
DB 307 CGCTAACAGCGAGACACTATAGAGTGTTCAGAGTTCTTACGAGCTCTGCAATC 248
QY 207 TGAGCCCGAGGGGCTGCGGCGGCTGACCGCGCGCGG 244
DB 247 GGAGCTGACAGACGTGCGCGCGCACCCGACGCGCAGAG 210

RESULT 12
US-09-252-991A-9783
Sequence 9783, Application US/09252991A
Patent No. 6551735
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO: 9783
LENGTH: 1959
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9783

Query Match 13.4%; Score 36.4; DB 4; Length 1959;
Best Local Similarity 51.9%; Pred. No. 2.1;
Matches 82; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 87 CACCTCCTGAGCCCTGCTACAGCTGTGCGGACATGTGGCCGACTTTCTGCGAGGAA 146
DB 1587 CAACCTCACTCTCTGAGCTGACCTTACCCGCACTGAGGTGCGCTTGACCGGAT 1646
QY 147 CCCCACTCAGTGTGAGGATTTCAACGCGCGCTTACTTCTGTGAGAACCGCAAGC 206
DB 1647 CGCTAACAGCGAGACACTATAGAGTGTTCAGAGTTCTTACGAGCTCTGCAATC 1706
QY 207 TGAGCCCGAGGGGCTGCGGCGGCTGACCGCGCGCGG 244
DB 1707 GGAGCTGACAGACGTGCGCGCGCACCCGACGCGCAGAG 1744

RESULT 13
US-09-105-537-34
Sequence 34, Application US/09105537A

Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/105.537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 34
LENGTH: 4689
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-105-537-34

Query Match 13.2%; Score 35.8; DB 3; Length 4689;
Best Local Similarity 46.9%; Pred. No. 3.3;
Matches 112; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 4 GGCTGCAGGTGGAATTGCTCTTCTTCCGCTACATCTGGAATGGAAGGACCTAGACCTTGC 63
DB 3406 GCCCACACGAGAGCCGACACCTCTCTCTGACGCGAGCGGAAACAGCCCGGAA 3465
QY 64 CGCTGTACCGGCTCACTGTCTTCACTTCTGAGACCCCTGTAGATGTGCGCGGACAT 123
DB 3466 GCCACCACTCACTCCGCACTCACTGATGAGGCGCGCGGTACATGCGCGCTGC 3525
QY 124 GTGCGGACCTTCTGCGAGGAAACCCCACTCACTGATGAGATTTCAACGCGCGCTC 183
DB 3526 GACGTGCGGACCCGACCCGACCCGACCTTCTTCAAGGATCTCCGCGAGAGCGCC 3585
QY 184 TACTTGTGAGAGCCGACGAGGCTGAGCCGAGGAGCTGCGGCGCTGACCGCGCGG 242
DB 3586 CTACCGCGGTGTTCACACCGCGCGGCGGAGATCGCTGAGCTGACCGG 3644

RESULT 14
US-09-105-537-5
Sequence 5, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/105.537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 5
LENGTH: 36778
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-105-537-5

Query Match 13.2%; Score 35.8; DB 3; Length 36778;
Best Local Similarity 46.9%; Pred. No. 4;
Matches 112; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 4 GGCTGCAGGTGGAATTGCTCTTCTTCCGCTACATCTGGAATGGAAGGACCTAGACCTTGC 63
DB 30396 GCCCACACGAGAGCCGACACCTCTCTCTGACGCGAGCGGAAACAGCCCGGAA 30455
QY 64 CGCTGTACCGGCTCACTGTCTTCACTTCTGAGACCCCTGTAGATGTGCGCGGACAT 123
DB 30456 GCCACCACTCACTCCGCACTCACTGAGGCGCGCGGTACCATGCGCGCTGC 30515
QY 124 GTGCGGACCTTCTGCGAGGAAACCCCACTCACTGAGATTTTACCGCGCGCTC 183

DB 30516 GACGTCGCCGACCCCGACGCGCATGCGCAACCTCTCGAGCCATCCCGCGAGACGCC 30575
QY 184 TACTTCTGTGAGGACCGGACGCTGAGCCCGAGGGCTGCGGGCTGCACCGCGCGG 242
DB 30576 CTCACCGCGCTGTCCACACCGCGCGCGCATCCGCGCGGATCCGCTGACGTCAACGG 30634

RESULT 15

US-09-320-878-19

; Sequence 19, Application US/09320878A

; Patent No. 6117659

; GENERAL INFORMATION:

; APPLICANT: ASHLEY, Gary

; APPLICANT: BETLACH, Mary C.

; APPLICANT: MCDANIEL, Robert

; APPLICANT: TANG, Li

; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE

; FILE REFERENCE: 300622002120

; CURRENT APPLICATION NUMBER: US/09/320,878A

; EARLIER FILING DATE: 1999-05-27

; EARLIER APPLICATION NUMBER: CIP OF 09/141,908

; EARLIER FILING DATE: 1998-08-28

; EARLIER APPLICATION NUMBER: CIP OF 09/073,538

; EARLIER FILING DATE: 1998-05-06

; EARLIER APPLICATION NUMBER: CIP OF 08/846,247

; EARLIER FILING DATE: 1997-04-30

; EARLIER APPLICATION NUMBER: 60/119,139

; EARLIER FILING DATE: 1999-02-08

; EARLIER APPLICATION NUMBER: 60/100,880

; EARLIER FILING DATE: 1998-09-22

; EARLIER APPLICATION NUMBER: 60/087,080

; EARLIER FILING DATE: 1998-05-28

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 19

; LENGTH: 38506

; TYPE: DNA

; ORGANISM: Streptomyces venezuelae

US-09-320-878-19

Query Match 13.2%; Score 35.8; DB 3; Length 38506;

Best Local Similarity 46.9%; Pred. No. 4;

Matches 112; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 4 GGCTGCGACGTGGAATTGCTTCTCCGCTACATCTGAGCTGGAGCTTACCTAGACCTGAC 63
DB 28538 GCTTCACGACGAGCCGACCACTCTCTCTCTGTCAGCCGACGCGAAGACCCCGGA 28597
QY 64 CGCTGTACCGCTCTCACTGTTTCATCTCTGAGCCCTGTACGACTGTGCGGACAT 123
DB 28598 GCCACCACTCAACCGCGACACTCAACGATGCGGCGCGGTACACATGCGCGCTGC 28657
QY 124 GTGCGCACTTTTGTGAGGAGAAACCCCACTCACTGAGATTTTCAACCGCGCTC 183
DB 28658 GACGTGCGGACCCCGACGCGCATGCGACCTCTGAGCCATCCCGCGAGAGCGCC 28717
QY 184 TACTTCTGTGAGGACCGGACGCTGAGCCCGAGGGCTGCGGGCTGCACCGCGCGG 242
DB 28718 CTCACCGCGCTGTCCACACCGCGCGGACCGCGCGCGATCCGCTGAGCGTCAACGG 28776

Search completed: March 13, 2004, 00:51:20
Job time: 26.1637 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 21:22:51 ; Search time 95.093 Seconds

(without alignments)
10491.276 Million cell updates/sec

Title: US-09-966-880a-13

Sequence: 1 aacggctgcacacgctggaatt.....agccatcatgacccctcaag 271

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 243257 seqs, 184079884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09C_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 271 | 100.0 | 271 | 9 | US-09-966-880a-13 |
| 2 | 271 | 100.0 | 2818 | 9 | US-09-966-880a-7 |
| 3 | 271 | 100.0 | 6564 | 9 | US-09-966-880a-10 |
| 4 | 271 | 100.0 | 11204 | 9 | US-09-966-880a-35 |
| 5 | 221.6 | 81.8 | 2440 | 9 | US-09-966-880a-1 |
| 6 | 87.4 | 32.3 | 2913 | 15 | US-10-027-632-112023 |
| 7 | 87.4 | 32.3 | 2913 | 15 | US-10-027-632-112024 |
| 8 | 87.4 | 32.3 | 201143 | 12 | US-10-240-425-1099 |
| 9 | 85.8 | 31.7 | 987 | 9 | US-09-925-300-699 |
| 10 | 85.8 | 31.7 | 1143 | 13 | US-10-098-841-268 |
| 11 | 85 | 31.4 | 1155 | 15 | US-10-460-923-1 |
| 12 | 85 | 31.4 | 1534 | 9 | US-09-729-674-173 |
| 13 | 85 | 31.4 | 2151 | 9 | US-09-822-830a-359 |
| 14 | 78.4 | 28.9 | 371 | 10 | US-09-918-995-30237 |
| 15 | 76.8 | 28.3 | 823 | 14 | US-10-247-671-62 |

| | | | | | | |
|----|------|------|---------|----|----------------------|--------------------|
| 16 | 76.8 | 28.3 | 1008 | 15 | US-10-131-827-8890 | Sequence 8890, Ap |
| 17 | 76.8 | 28.3 | 1348 | 9 | US-09-880-107-3282 | Sequence 3282, Ap |
| 18 | 68.4 | 25.2 | 456 | 14 | US-10-029-386-25187 | Sequence 25187, A |
| 19 | 68.4 | 25.2 | 1404 | 14 | US-10-198-846-13472 | Sequence 13472, A |
| 20 | 68.4 | 25.2 | 2773 | 15 | US-10-104-047-1759 | Sequence 1759, Ap |
| 21 | 64.8 | 23.9 | 572 | 9 | US-09-864-761-9416 | Sequence 9416, Ap |
| 22 | 63 | 23.2 | 476 | 10 | US-09-918-995-32385 | Sequence 32385, A |
| 23 | 50.6 | 18.7 | 445 | 15 | US-10-378-029-25 | Sequence 25, Appl |
| 24 | 50.6 | 18.7 | 879 | 14 | US-10-157-031-13 | Sequence 13, Appl |
| 25 | 49.4 | 18.2 | 556 | 9 | US-09-796-692-5061 | Sequence 5061, Ap |
| 26 | 49.4 | 18.2 | 556 | 14 | US-10-040-862-5061 | Sequence 5061, Ap |
| 27 | 49.4 | 18.2 | 556 | 15 | US-10-057-475B-5061 | Sequence 5061, Ap |
| 28 | 49.4 | 18.2 | 556 | 15 | US-10-154-884B-5061 | Sequence 5061, Ap |
| 29 | 48 | 17.7 | 553 | 9 | US-09-796-692-4504 | Sequence 4504, Ap |
| 30 | 48 | 17.7 | 553 | 14 | US-10-040-862-4504 | Sequence 4504, Ap |
| 31 | 48 | 17.7 | 553 | 15 | US-10-057-475B-4504 | Sequence 4504, Ap |
| 32 | 48 | 17.7 | 553 | 15 | US-10-154-884B-4504 | Sequence 4504, Ap |
| 33 | 41 | 15.1 | 821 | 15 | US-10-388-934-189 | Sequence 189, Ap |
| 34 | 40.6 | 15.0 | 807 | 12 | US-10-424-599-135749 | Sequence 135749, A |
| 35 | 38.4 | 14.2 | 922 | 12 | US-10-425-114-35403 | Sequence 35403, A |
| 36 | 38.4 | 14.2 | 1362 | 12 | US-10-425-114-35704 | Sequence 35704, A |
| 37 | 38.2 | 14.1 | 60 | 10 | US-09-908-978-7319 | Sequence 7319, Ap |
| 38 | 38 | 14.0 | 1929 | 10 | US-09-899-642-1 | Sequence 1, Appl1 |
| 39 | 37.2 | 13.7 | 1323 | 14 | US-10-156-761-1176 | Sequence 1176, Ap |
| 40 | 37.2 | 13.7 | 9025608 | 14 | US-10-156-761-1 | Sequence 1, Appl1 |
| 41 | 37 | 13.7 | 520 | 14 | US-10-184-634-332 | Sequence 332, App |
| 42 | 37 | 13.7 | 520 | 14 | US-10-184-634-332 | Sequence 332, App |
| 43 | 37 | 13.7 | 674 | 14 | US-10-257-826A-150 | Sequence 150, App |
| 44 | 37 | 13.7 | 1503 | 14 | US-10-156-761-5170 | Sequence 5170, Ap |
| 45 | 37 | 13.7 | 9025608 | 14 | US-10-156-761-1 | Sequence 1, Appl1 |

ALIGNMENTS

| | |
|--|--|
| RESULT 1 | |
| US-09-966-880a-13 | |
| Sequence 13, Application US/09966880A | |
| Patent No. US2002016473A1 | |
| GENERAL INFORMATION: | |
| APPLICANT: Honto, Tasuku | |
| APPLICANT: Muramatsu, Masamichi | |
| TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE | |
| FILE REFERENCE: 06501-088001 | |
| CURRENT APPLICATION NUMBER: US/09/966, 880A | |
| CURRENT FILING DATE: 2001-09-28 | |
| PRIOR APPLICATION NUMBER: PCT/JP00/01918 | |
| PRIOR FILING DATE: 2000-03-28 | |
| PRIOR APPLICATION NUMBER: JP 11-371382 | |
| PRIOR FILING DATE: 1999-12-27 | |
| PRIOR APPLICATION NUMBER: JP 11-178999 | |
| PRIOR FILING DATE: 1999-06-24 | |
| PRIOR APPLICATION NUMBER: JP 11-87192 | |
| PRIOR FILING DATE: 1999-03-29 | |
| NUMBER OF SEQ ID NOS: 36 | |
| SOFTWARE: FastSeq for Windows Version 4.0 | |
| SEQ ID NO 13 | |
| LENGTH: 271 | |
| TYPE: DNA | |
| ORGANISM: Homo sapiens | |
| US-09-966-880a-13 | |
| Query Match | |
| Best Local Similarity 100.0%; Score 271; DB 9; Length 271; | |
| Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| Cy | |
| 1 AACGGCTGCCACGATGATGCTCTTCTCCGCTACATCTCGAGCTGGAGCTAGACCT 60 | |
| Db | |
| 1 AACGGCTGCCACGATGATGCTCTTCTCCGCTACATCTCGAGCTGGAGCTAGACCT 60 | |
| Cy | |
| 61 GAGCGCTGCTACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120 | |
| Db | |
| 61 GAGCGCTGCTACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120 | |

```
QY 121 CATGTCGCGCACTTTCTGTGAGGAGAACCCCAACTCAGTCTGAGGATCTTTCACCGCGGC 180
Db 121 CATGTCGCGCACTTTCTGTGAGGAGAACCCCAACTCAGTCTGAGGATCTTTCACCGCGGC 180
QY 181 CTCTACTTCTGTGAGGAGCAAGGCTGAGCCCGAGGAGGCTGCGCGCTGACCGCGCC 240
Db 181 CTCTACTTCTGTGAGGAGCAAGGCTGAGCCCGAGGAGGCTGCGCGCTGACCGCGCC 240
QY 241 GGGGTGCAAAATAGCCATCATGACCTTCAAG 271
Db 241 GGGGTGCAAAATAGCCATCATGACCTTCAAG 271

RESULT 2
US-09-966-880A-7
; Sequence 7, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2818
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (80)...(673)
; FEATURE:
; NAME/KEY: 5' UTR
; LOCATION: (1)...(79)
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: (677)...(2818)
; US-09-966-880A-7

Query Match 100.0%; Score 271; DB 9; Length 2818;
Best Local Similarity 100.0%; Pred. No. 1.9e-68;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AACGCTGCCAGCTGAATGCTCTTCTCCGCTATCATCTCGAGCTGGAGCTTAACCT 60
Db 236 AACGCTGCCAGCTGAATGCTCTTCTCCGCTATCATCTCGAGCTGGAGCTTAACCT 295
QY 61 GGCCTGCTGACCGGCTGACCTGTTCACTCTGTGAGGCCCTGCTACAGACTGTGCCGA 120
Db 296 GGCCTGCTGACCGGCTGACCTGTTCACTCTGTGAGGCCCTGCTACAGACTGTGCCGA 355
QY 121 CATGTGGCGCACTTTCTGCGAGGGAAGCCCAACTCAGTCTGAGATCTTTCACCGCGCC 180
Db 356 CATGTGGCGCACTTTCTGCGAGGGAAGCCCAACTCAGTCTGAGATCTTTCACCGCGCC 415
QY 181 CTCTACTTCTGTGAGGAGCAAGGCTGAGCCCGAGGAGGCTGCGCGCTGACCGCGCC 240
Db 416 CTCTACTTCTGTGAGGAGCAAGGCTGAGCCCGAGGAGGCTGCGCGCTGACCGCGCC 475
QY 241 GGGGTGCAAAATAGCCATCATGACCTTCAAG 271
Db 476 GGGGTGCAAAATAGCCATCATGACCTTCAAG 506

RESULT 3
US-09-966-880A-10
; Sequence 10, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 6564
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-966-880A-10

Query Match 100.0%; Score 271; DB 9; Length 6564;
Best Local Similarity 100.0%; Pred. No. 1.9e-68;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 35
LENGTH: 11204
TYPE: DNA
ORGANISM: Homo sapiens
US-09-966-880A-35

Query Match 100.0%; Score 271; DB 9; Length 11204;
Best Local Similarity 100.0%; Pred. No. 2e-68;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGCTCCAGCTGGAATGCTCTCTCCCTGCTACATCTGAGACTGGAGACTAGACCT 60
DB 7807 AACGGCTCCAGCTGGAATGCTCTCTCCCTGCTACATCTGAGACTGGAGACTAGACCT 7866
QY 61 GAGCGCTGCTACCGCGTCACTGTTCACTCTGAGCCCTGCTACGACTGTGCCGA 120
DB 7867 GAGCGCTGCTACCGCGTCACTGTTCACTCTGAGCCCTGCTACGACTGTGCCGA 7926
QY 121 CATGTGCGGACTTCTGCGAGGGAACCCCACTCACTGAGAGATTTCAACCGCGC 180
DB 7927 CATGTGCGGACTTCTGCGAGGGAACCCCACTCACTGAGAGATTTCAACCGCGC 7986
QY 181 CTCTACTTCTGTGAGAGACCGCAAGGCTGAGCCGCGCTGCGCGCTGCAACCGCGC 240
DB 7987 CTCTACTTCTGTGAGAGACCGCAAGGCTGAGCCGCGCTGCGCGCTGCAACCGCGC 8046
QY 241 GGGGTGCAATPAGCATCATGACCTTCAAG 271
DB 8047 GGGGTGCAATPAGCATCATGACCTTCAAG 8077

RESULT 5

US-09-966-880A-1
Sequence 1, Application US/09966880A
Patent No. US2002016743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966.880A
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2440
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (93)...(686)
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)...(92)
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (690)...(2440)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(2440)
OTHER INFORMATION: n = A,T,C or G
US-09-966-880A-1

Query Match 81.8%; Score 221.6; DB 9; Length 2440;
Best Local Similarity 89.2%; Pred. No. 3e-54;
Matches 239; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 4 GGGTGGCAAGGAAATGCTCTTCTCCGCTACATCTGAGACTGGAGACTTAGACCTTGGC 63
DB 252 GGGTGGCAAGGAAATGCTCTTCTCCGCTACATCTGAGACTGGAGACTTAGACCTTGGC 311
QY 64 CGGTGTACCGGCTACCGCTGTTCACTCCCTGAGAGCCCTGCTAGACATGAGCCGACAT 123
DB 312 CGGTGTACCGGCTACCGCTGTTCACTCCCTGAGAGCCCTGCTAGACATGAGCCGACAT 371
QY 124 GTGGCCGACTTCTGCGAGGGAACCCCACTCACTGAGATCTTCAACCGCGCCTC 183
DB 372 GTGGCCGACTTCTGCGAGGGAACCCCACTCACTGAGATCTTCAACCGCGCCTC 431
QY 184 TACTTGTGAGAGACCGAAGGCTGAGCCCGAGGAGCTGCGGAGCTGACCGCGCGG 243
DB 432 TACTTGTGAGAGACCGAAGGCTGAGCCCGAGGAGCTGCGGAGCTGACCGCGCGG 491
QY 244 GTGCAATPAGCATCATGACCTTCAAG 271
DB 492 GTGCAATPAGCATCATGACCTTCAAG 519

RESULT 6

US-10-027-632-112023
Sequence 112023, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Mang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
POLYMORPHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027.632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 112023
LENGTH: 2913
TYPE: DNA
ORGANISM: Human
US-10-027-632-112023

Query Match 32.3%; Score 87.4; DB 15; Length 2913;
Best Local Similarity 60.8%; Pred. No. 1.2e-15;
Matches 161; Conservative 0; Mismatches 101; Indels 3; Gaps 1;

QY 7 TGGCAGGTGAATGCTCTTCTCCGCTACATCTGAGACTGGAGACTTAGACCTTGGCCGC 66
DB 11 TGGCAGGTGAATGCTCTTCTCCGCTACATCTGAGACTGGAGACTTAGACCTTGGCCGC 70
QY 67 TGGTACCGGCTACCGCTGTTCACTCTGAGAGCCCTGCTAGAGCTGAGCCGACATGTG 126
DB 71 TGGTACCGGCTACCGCTGTTCACTCTGAGAGCCCTGCTAGAGCTGAGCCGACATGTG 130
QY 127 GCGGACTTCTGCGAGGGAACCCCACTCACTGAGATCTTCAACCGCGGCGCTTAC 186
DB 131 GCGGACTTCTGCGAGGGAACCCCACTCACTGAGATCTTCAACCGCGGCGCTTAC 190

QY 187 TTCTGTGAGGACCGGAGGCTGAGCCGAGGGGCTGGGGGGCTGACACCGGCGGGGGT 246
DB 191 TACTTCTGGGA--TACAGATTACAGAGGGGCTCCGACCTGAGTCAAGAGGGGCC 247
QY 247 CAATAGCCATCATGACCTTCAAG 271
DB 248 TCCGTGAGATCATGGGCTACAAAG 272

RESULT 7
US-10-027-632-112024
Sequence 112024, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
POLYMORPHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108827.129
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 112024
LENGTH: 2913
TYPE: DNA
ORGANISM: Human
US-10-027-632-112024

Query Match 32.3%; Score 87.4; DB 15; Length 2913;
Best Local Similarity 60.8%; Pred. No. 1.2e-15;
Matches 161; Conservative 0; Mismatches 101; Indels 3; Gaps 1;

QY 7 TCCACGTTGGAATTGCTCTTCTCCGCTACATCTGGAGCTGGAGCTGACCTGGAGCCG 66
DB 11 TGTTCATGACGAAGGCTCTTCTCTTGTGTGACGACATCTCTCTTAACACA 70
QY 67 TGTACCGGCTACCGTGTTCACCTCTGGAGCCCTGTACGACTGTGCGGACATGTG 126
DB 71 AACTAGAGGTACCGTGTACATCTTGGAGCCCTTGGCAGATGTGAGGGAGGTG 130
QY 127 GCCGACTTCTGCGAGGAGAACCCCACTCACTGTGAGAGATTTGACCGGCGCTTAC 186
DB 131 GCCGAGTCTGCGGAGGAGACGAGACGATCTCAACATTTACCGCGCGCTTAC 190
QY 187 TTCTGTGAGGACCGCAAGGCTGAGCCGAGGGGCTGGCGGCTGACCGCGCGGGT 246
DB 191 TACTTCTGGGA--TACAGATTACAGAGGGGCTCCGAGCTGAGTCAAGAGGGGCC 247
QY 247 CAATAGCCATCATGACCTTCAAG 271
DB 248 TCCGTGAGATCATGGGCTACAAAG 272

RESULT 8
US-10-240-425-1099
Sequence 1099, Application US/10240425
Publication No. US2004003502A1
GENERAL INFORMATION:

APPLICANT: Williams, Amanda
APPLICANT: Boland, Joseph F.
APPLICANT: Lord, Reginald V.
APPLICANT: Alvarez, Chris
APPLICANT: Wetzel, Jon C.
APPLICANT: Schert, Uwe
APPLICANT: Vockley, Joseph G.
TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
FILE REFERENCE: 44921-5026
CURRENT APPLICATION NUMBER: US/10/240,425
CURRENT FILING DATE: 2002-09-30
PRIOR APPLICATION NUMBER: PCT/US01/09847
PRIOR FILING DATE: 2001-03-28
PRIOR APPLICATION NUMBER: US 60/193,446
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 1588
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1099
LENGTH: 201143
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US2004003502A1 AL022318
US-10-240-425-1099

Query Match 32.3%; Score 87.4; DB 12; Length 201143;
Best Local Similarity 60.8%; Pred. No. 1.4e-15;
Matches 161; Conservative 0; Mismatches 101; Indels 3; Gaps 1;

QY 7 TCCACGTTGGAATTGCTCTTCTCCGCTACATCTGGAGCTGGAGCTGACCTGGAGCCG 66
DB 167899 TGTTCATGACGAAGGCTCTTCTCTTGTGTGACGACATCTCTCTTAACACA 167958
QY 67 TGTACCGGCTACCGTGTTCACCTCTGGAGCCCTGTACGACTGTGCGGACATGTG 126
DB 167959 AACTAGAGGTACCGTGTACATCTTGGAGCCCTTGGCAGATGTGAGGGAGGTG 168018
QY 127 GCCGACTTCTGCGAGGAGAACCCCACTCACTGTGAGAGATTTGACCGGCGCTTAC 186
DB 168019 GCCGAGTCTGCGGAGGAGACGAGACGATCTCAACATTTACCGCGCGCTTAC 168078
QY 187 TTCTGTGAGGACCGCAAGGCTGAGCCGAGGGGCTGGCGGCTGACCGCGCGGGT 246
DB 168079 TACTTCTGGGA--TACAGATTACAGAGGGGCTCCGAGCTGAGTCAAGAGGGGCC 168135
QY 247 CAATAGCCATCATGACCTTCAAG 271
DB 168136 TCCGTGAGATCATGGGCTACAAAG 168160

RESULT 9
US-09-925-300-699
Sequence 699, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
APPLICANT: Steve Ruben,
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 699
LENGTH: 987
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-300-699

Query Match 31.7%; Score 85.8; DB 9; Length 987;
Best Local Similarity 60.4%; Pred. No. 3.4e-15;
Matches 160; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

QY 7 TCCACGCGGAATTCCTCTTCTCCGCTACATCTGGAGTGGAGACCTTGGCCGC 66
DB 291 TGTCAATGCAAGAGGCTTCTCTTGTGTTCTGCAAGACATCTCTCTTAACA 350
QY 67 TGTACCGGCTACCTGCTTCACTTCTGGAGCCCTGTAGATTTGGCCGACATGTG 126
DB 351 AAGTACAGAGTACCTGCTGACATCTTGGAGCCCTTGCAGACTGTGCAAGGAGTG 410
QY 127 GCCGCTTCTGCAAGGGAACCCCACTCAGTCTGAGATTTTACCGCGGCTCTAC 186
DB 411 GCCGAGTCTTCCGCGGACGACAGCAAGTGAATTTCAACATTTTACCGCCGCTTAC 470
QY 187 TTCTGTAGAGACCGCAAGGCTGAGCCCGAGGCTGCGGCGCTGACCGCGCGGAGTG 246
DB 471 TACTTCAAGTATC---CATGTTACAGAGAGGGGCTCCGAGCCTGAGTCAAGAGGGGTC 527
QY 247 CAATAGCCATCATGACTTCAAG 271
DB 528 GCTGTGAGATCATGACTATGAAG 552

RESULT 10

US-10-098-841-268
Sequence 268, Application US/10098841
Publication No. US20020197679A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Xu, Chongjun
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Zhao, Jian-Rui
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: Weinman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Qian, Xiaohong B.
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/10/098,841
CURRENT FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 331
SOFTWARE: PC_FL_genes Version 1.0
SEQ ID NO 268
LENGTH: 1143
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (128)..(700)
US-10-098-841-268

Query Match 31.7%; Score 85.8; DB 13; Length 1143;
Best Local Similarity 60.4%; Pred. No. 3.4e-15;
Matches 160; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

QY 7 TGCACGTGAAATTCCTCTTCTCCGCTACATCTCGAGTGGAGCTTAGACCTTGGCCGC 66

DB 320 TGTATGCAAGAAAGGCTTCTCTCTGTTCTGCAAGACATATGTTCTCTTAACA 379
QY 67 TGTACCGGCTACCTGCTTCACTTCTGGAGCCCTGTAGACTGTGCGGACATGTG 126
DB 380 AAGTACAGAGTACCTGCTGACATTTTGAAGCCCTTGCACAGACTGTGCAAGGAGTG 439
QY 127 GCCGCTTCTGCAAGGGAACCCCACTCAGTCTGAGATTTTCAACCGCGGCTCTAC 186
DB 440 GCCGAGTCTTCCGCGGACGACAGCAAGTGAATTTCAACATTTTACCGCCGCTTAC 499
QY 187 TTCTGTAGAGACCGCAAGGCTGAGCCCGAGGCTGCGGCGCTGCAACGCGCGGAGTG 246
DB 500 TACTTCAAGTATC---CATGTTACAGAGAGGGGCTCCGAGCCTGAGTCAAGAGGGGTC 556
QY 247 CAATAGCCATCATGACTTCAAG 271
DB 557 GCTGTGAGATCATGACTATGAAG 581

RESULT 11

US-10-460-923-1
Sequence 1, Application US/10460923
Publication No. US2004000951A1
GENERAL INFORMATION:
APPLICANT: MALIM, Michael H.
APPLICANT: SHEEHY, Ann M.
APPLICANT: HARRIS, Reuben S.
APPLICANT: BISHOP, Kate N.
APPLICANT: NEUBERGER, Michael S.
APPLICANT: GADDIS, Nathan C.
APPLICANT: SIMON, James H.M.
TITLE OF INVENTION: DNA Deamination Mediates Innate Immunity to Retroviral Infection
FILE REFERENCE: 22253-74380
CURRENT APPLICATION NUMBER: US/10/460,923
CURRENT FILING DATE: 2003-06-13
PRIOR APPLICATION NUMBER: US 60/388,513
PRIOR FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: US 60/472,952
PRIOR FILING DATE: 2003-05-23
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1155
TYPE: DNA
ORGANISM: Homo sapiens
US-10-460-923-1

Query Match 31.4%; Score 85; DB 15; Length 1155;
Best Local Similarity 60.3%; Pred. No. 5.7e-15;
Matches 161; Conservative 0; Mismatches 100; Indels 6; Gaps 1;

QY 2 ACGGCTGCAACGCGGAATTCCTCTTCTCCGCTACATCTGAGTGGAGACTTGAACCTG 61
DB 761 AAGCGCGCATGCAAGACTGTCTTCTGAGCTGATTCCTTTGGAAGCTGAGACTG 820
QY 62 GCCGCTGCTACCGGCTGACCTGCTTCACTTCTGAGGCCCCCTGCTAGACTGTGCGGAC 121
DB 821 ACCAGAGTACAGGCTTACTGCTTACCTTCCGAGAGCCCTGCTTCACTGAGTGGCCAG 880
QY 122 ATGTGCGGACTTTTCTGCGAGGGAACCCCACTCAGTCTGAGAGTCTTACCGCGGCC 181
DB 881 AATGCTTAATTCATTTCAAAAAACAAACACGTGAGCTGTGCATCTTACTGCCGCA 940
QY 182 TCTACTTCTGTGAGACCGGAGAGGCTGAGCCGAGGGGCTGCGGCGCTGCAACGCGCG 241
DB 941 TCTA-----TGATGATCAAGAAAGATGTAGAGAGGGCTGCGCACTTGGAGGCTG 994
QY 242 GGTGCAATAGCATCATGACTTCA 268
DB 995 GGCGCAAAATTTCAATAATGACATACA 1021

RESULT 12

US-09-729-674-173
Sequence 173, Application US/09729674
Patent No. US20010039335A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John W.
APPLICANT: Lavallee, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Agostino, Michael J.
APPLICANT: Steindinger II, Robert J.
APPLICANT: Spaulding, Vikki
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Reichel, Kim
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: 6055-64X
CURRENT APPLICATION NUMBER: US/09/729,674
CURRENT FILING DATE: 2000-12-04
PRIOR APPLICATION NUMBER: 09/539,330
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 283
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 173
LENGTH: 1534
TYPE: DNA
ORGANISM: Homo sapiens
US-09-729-674-173

Query Match 31.4%; Score 85; DB 9; Length 1534;
Best Local Similarity 60.3%; Pred. No. 5,8e-15;
Matches 161; Conservative 0; Mismatches 100; Indels 6; Gaps 1;

QY 2 ACGGCTGCACGATGGAATGCTCTCCCTGCTACATCTCGAGAGGAGACCTAGACCTG 61
DB 811 AAGGCCCGCATGACAGAGCTGTCTTCCTGAGACGTATTCCTTTGGAAGCTGAGACCTGG 870
QY 62 GCCGCTGTACCGCGCTACCTGTTACCTCCTCGAGGCCCTGCTACGACTGTGCGCCGAC 121
DB 871 ACCAGACTACAGGGTTACTCTTCACTCCTCGAGGCCCTGCTACGCTGTGCGCCAG 930
QY 122 ATGTGGCCGACTTTCTGCGAGGAAACCCCACTGATGTGAGATCTTCAACCGGCGCC 181
DB 931 AATGGCTAATTCATTTCAAAAACAAACAGTGAAGCTGTGATCTTCACTGCGCCGA 990
QY 182 TCTACTCTGTGAGACCGCAAGCGCTGAGCCCGAGGGGCTGGGGGCTGACACCGCGCG 241
DB 991 TCTA-----TGATGATCAAGGAAGATGTGAGAGGGGCTGGCAACCTGGCCGAGCTG 1044
QY 242 GGGTGAATAGCATCATGACCTTCA 268
DB 1045 GGGCCAAATTTCAATATGACATACA 1071

RESULT 13

US-09-822-830A-359
Sequence 359, Application US/09822830A
Patent No. US20020142952A1
GENERAL INFORMATION:
APPLICANT: Genetics Institute, Inc.
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Reichel, Kim
APPLICANT: Agostino, Michael J.
APPLICANT: Howes, Steven H.
APPLICANT: Resnick, Richard J.
APPLICANT: Gulukota, Kamalakhar
APPLICANT: Graham, James R.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS

FILE REFERENCE: GIN 6402

CURRENT APPLICATION NUMBER: US/09/822,830A
CURRENT FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/195,604
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 631
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 359
LENGTH: 2151
TYPE: DNA
ORGANISM: Homo sapiens
US-09-822-830A-359

Query Match 31.4%; Score 85; DB 9; Length 2151;
Best Local Similarity 60.3%; Pred. No. 5,9e-15;
Matches 161; Conservative 0; Mismatches 100; Indels 6; Gaps 1;

QY 2 ACGGCTGCACGATGGAATGCTCTCCCTGCTACATCTCGAGAGGAGCTAGACCTAGACCTG 61
DB 1450 AAGGCCCGCATGACAGAGCTGTCTTCCTGAGACGTATTCCTTTGGAAGCTGAGACTGG 1509
QY 62 GCCGCTGTACCGCGCTACCTGTTACCTCCTCGAGGCCCTGCTACGACTGTGCGCCGAC 121
DB 1510 ACCAGACTACAGGGTTACTCTTCACTCCTCGAGGCCCTGCTACGCTGTGCGCCAG 1569
QY 122 ATGTGGCCGACTTTCTGCGAGGAAACCCCACTGATGTGAGATCTTCAACCGGCGCC 181
DB 1570 AATGGCTAATTCATTTCAAAAACAAACAGTGAAGCTGTGATCTTCACTGCGCCGA 1629
QY 182 TCTACTCTGTGAGACCGCAAGCGCTGAGCCCGAGGGGCTGGGGGCTGACACCGCGCG 241
DB 1630 TCTA-----TGATGATCAAGGAAGATGTGAGAGGGGCTGGCAACCTGGCCGAGGCTG 1683
QY 242 GGGTGAATAGCATCATGACCTTCA 268
DB 1684 GGGCCAAATTTCAATATGACATACA 1710

RESULT 14

US-09-918-995-30237
Sequence 30237, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 30237
LENGTH: 371
TYPE: DNA
ORGANISM: Homo sapiens
US-09-918-995-30237

Query Match 28.9%; Score 78.4; DB 10; Length 371;
Best Local Similarity 64.1%; Pred. No. 4,4e-13;
Matches 118; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 2 ACGGCTGCACGATGGAATGCTCTCCCTGCTACATCTCGAGAGGAGCTAGACCTAGACCTG 61
DB 171 AAGGCCCGCATGACAGAGCTGTCTTCCTGAGACGTATTCCTTTGGAAGCTGAGACTGG 230
QY 62 GCCGCTGTACCGCGCTACCTGTTACCTCCTCGAGGCCCTGCTACGACTGTGCGCCGAC 121
DB 231 ACCAGACTACAGGGTTACTCTTCACTCCTCGAGGCCCTGCTACGCTGTGCGCCAG 290
QY 122 ATGTGGCCGACTTTCTGCGAGGAAACCCCACTGATGTGAGATCTTCAACCGGCGCC 181

Db 291 AAATGCTAAATTCATTTCAAAAAACAACAGTGAGCTGTGATCTTCACTGCGCGCA 350
QY 182 TCTA 185
Db 351 TCTA 354

RESULT 15

US-10-247-671-62
; Sequence 62, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 62
; LENGTH: 823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 135626CB1
US-10-247-671-62

Query Match 28.3%; Score 76.8; DB 14; Length 823;
Best Local Similarity 60.5%; Pred. No. 1.3e-12;
Matches 167; Conservative 0; Mismatches 97; Indels 12; Gaps 2;

QY 2 ACCGCTGCCACGTGATGCTCTTCTCCGCTACATCTGAGACTGGAGACCTAGACCTG 61
Db 115 ACCGCTGCCACGTGAGCTGCTCTTCTCCGCTACATCTGAGACTGGAGACCTAGACCTG 174
QY 62 GCGGCTGCTACCGCGTCACTGCTGAGAGCCCGTACTAC-----GACTG 115
Db 175 CCGAGTCTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 234
QY 116 CCGACATGTGGCGGCTTCTGCGAGGAGACCCCACTCACTGAGATCTTCAACG 175
Db 235 CCGGGAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 294
QY 176 CCGGCTCTACTTCTGAGAGACCGCAAGGCTGAGCCCGAGGCGCTGCGCGCTGCA 235
Db 295 CCGCATCTATGATGATGACCC-----TATATAGAGGCGCTGCAATGCTGCGG 348
QY 236 GCGCCGCGGTGCAATAGCATGATGATGATGATGATGATGATGATGATGATGATG 271
Db 349 ATGCTGGGCGCCAGTCTCATCATGATGATGATGATGATGATGATGATGATGATG 384

Search completed: March 13, 2004, 05:31:06
Job time: 105.099 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 13:19:34 ; Search time 862.933 Seconds

(without alignments)
9378.080 Million cell updates/sec

Title: US-09-966-880a-13

Perfect score: 271
Sequence: 1 aacggctgccacgtgaattc.....agccatcatgaccttcaag 271

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_espba:*
2: em_estum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estin:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_din:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rnd:*
26: em_gss_pig:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 271 | 100.0 | 541 | 10 | BF238155 601811880 |
| 2 | 271 | 100.0 | 693 | 12 | BG757392 602711022 |
| 3 | 271 | 100.0 | 743 | 12 | BG686133 602638412 |
| 4 | 271 | 100.0 | 820 | 12 | BG757089 602715124 |

| 5 | 271 | 100.0 | 853 | 13 | BK464579 |
|----|-------|-------|------|----|----------|
| 6 | 271 | 100.0 | 872 | 12 | BG758510 |
| 7 | 271 | 100.0 | 889 | 12 | BG686876 |
| 8 | 271 | 100.0 | 953 | 13 | BQ065440 |
| 9 | 271 | 100.0 | 1052 | 13 | BQ055935 |
| 10 | 271 | 100.0 | 1201 | 9 | AL559877 |
| 11 | 271 | 100.0 | 1201 | 9 | AL559877 |
| 12 | 269.4 | 99.4 | 942 | 13 | BK402063 |
| 13 | 257.4 | 95.0 | 535 | 14 | BF975166 |
| 14 | 221.6 | 81.8 | 522 | 12 | CD707143 |
| 15 | 221.6 | 81.8 | 653 | 10 | BG44705 |
| 16 | 221.6 | 81.8 | 1282 | 11 | BK637360 |
| 17 | 203 | 74.9 | 653 | 10 | AK080144 |
| 18 | 196 | 72.3 | 650 | 9 | BF975096 |
| 19 | 196 | 72.3 | 650 | 9 | AJ449745 |
| 20 | 196 | 72.3 | 696 | 9 | AJ450317 |
| 21 | 196 | 72.3 | 729 | 9 | AJ453647 |
| 22 | 196 | 72.3 | 757 | 9 | AJ450296 |
| 23 | 183.6 | 67.7 | 623 | 9 | AJ446140 |
| 24 | 118.8 | 43.8 | 632 | 14 | AJ450295 |
| 25 | 118.6 | 43.8 | 197 | 14 | CD471262 |
| 26 | 118.6 | 43.8 | 197 | 14 | CD826686 |
| 27 | 114 | 42.1 | 544 | 14 | CF722800 |
| 28 | 114 | 42.1 | 549 | 14 | CD472096 |
| 29 | 114 | 42.1 | 555 | 14 | CD469966 |
| 30 | 114 | 42.1 | 561 | 14 | CD470596 |
| 31 | 114 | 42.1 | 562 | 14 | CD471717 |
| 32 | 114 | 42.1 | 562 | 14 | CD470721 |
| 33 | 114 | 42.1 | 569 | 14 | CD471750 |
| 34 | 114 | 42.1 | 578 | 14 | CD471811 |
| 35 | 114 | 42.1 | 582 | 14 | CD472071 |
| 36 | 114 | 42.1 | 596 | 14 | CD471678 |
| 37 | 114 | 42.1 | 597 | 14 | CD470940 |
| 38 | 114 | 42.1 | 630 | 14 | CD469721 |
| 39 | 114 | 42.1 | 637 | 14 | CD470714 |
| 40 | 114 | 42.1 | 650 | 14 | CD471162 |
| 41 | 112.4 | 41.5 | 707 | 14 | CD470713 |
| 42 | 112.4 | 41.5 | 582 | 14 | CD469582 |
| 43 | 112.4 | 41.5 | 643 | 14 | CD465467 |
| 44 | 112.4 | 41.5 | 644 | 14 | CD469636 |
| 45 | 110.8 | 40.9 | 647 | 14 | CD469225 |
| | | | | | CD470661 |

ALIGNMENTS

RESULT 1
BF238155
LOCUS
DEFINITION
601811880F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4054915 5', mRNA sequence.
ACCESSION
BF238155
VERSION
BF238155.1 GI:11152074
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
NIH-MGC <http://mgc.nci.nih.gov/>
1 (Bases 1 to 541)
TITLE
Unpublished (1999)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgabs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LHC695 row: p column: 20
High quality sequence stop: 541.

FEATURES

source

Location/Qualifiers

1..541
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4054915"
/tissue_type="Primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH MGC 48"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 271; DB 10; Length 541;
Best Local Similarity 100.0%; Pred. No. 2.5e-47;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGCTGCACGTGGAATTGCTTCCCTCCGCTACATCTGGACTGGACCTAGACCT 60
DB 234 AACGGCTGCACGTGGAATTGCTTCCCTCCGCTACATCTGGACTGGACCTAGACCT 293
QY 61 GGCCGCTGCTACCGGCTACCTGCTTCACTCTGGAGCCCTGCTACGACTGTGCCGA 120
DB 294 GGCCGCTGCTACCGGCTACCTGCTTCACTCTGGAGCCCTGCTACGACTGTGCCGA 353
QY 121 CATGTGGCGGACCTTTCTGCGAGGAAACCCCAACTCACTGAGGATCTTCAACCGCGGC 180
DB 354 CATGTGGCGGACCTTTCTGCGAGGAAACCCCAACTCACTGAGGATCTTCAACCGCGGC 413
QY 181 CTCTACTTCTGTGAGGACCGCAAGGCTAGCCCGAGGGGCTGCGCGGCTGACCGCGCC 240
DB 414 CTCTACTTCTGTGAGGACCGCAAGGCTAGCCCGAGGGGCTGCGCGGCTGACCGCGCC 473
QY 241 GGGGTGCAATATGACCATCATGACCTTCAAG 271
DB 474 GGGGTGCAATATGACCATCATGACCTTCAAG 504

RESULT 2
BG757392 693 bp mRNA linear EST 15-MAY-2001
LOCUS 602711022F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4851580 5',
DEFINITION mRNA sequence.
ACCESSION BG757392
VERSION BG757392.1 GI:14068045
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 693)
NIH-MGC http://imgc.ncbi.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LLCM1634 row: k column: 05
High quality sequence stop: 693.
Location/Qualifiers

FEATURES

source

1..693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4851580"
/tissue_type="Primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH MGC 48"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 271; DB 12; Length 693;
Best Local Similarity 100.0%; Pred. No. 2.7e-47;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGCTGCACGTGGAATTGCTTCCCTCCGCTACATCTGGACTGGACCTAGACCT 60
DB 238 AACGGCTGCACGTGGAATTGCTTCCCTCCGCTACATCTGGACTGGACCTAGACCT 297
QY 61 GGCCGCTGCTACCGGCTACCTGCTTCACTCTGGAGCCCTGCTACGACTGTGCCGA 120
DB 298 GGCCGCTGCTACCGGCTACCTGCTTCACTCTGGAGCCCTGCTACGACTGTGCCGA 357
QY 121 CATGTGGCGGACCTTTCTGCGAGGAAACCCCAACTCACTGAGGATCTTCAACCGCGGC 180
DB 358 CATGTGGCGGACCTTTCTGCGAGGAAACCCCAACTCACTGAGGATCTTCAACCGCGGC 417
QY 181 CTCTACTTCTGTGAGGACCGCAAGGCTAGCCCGAGGGGCTGCGCGGCTGACCGCGCC 240
DB 418 CTCTACTTCTGTGAGGACCGCAAGGCTAGCCCGAGGGGCTGCGCGGCTGACCGCGCC 477
QY 241 GGGGTGCAATATGACCATCATGACCTTCAAG 271
DB 478 GGGGTGCAATATGACCATCATGACCTTCAAG 508

RESULT 3
BG686133 743 bp mRNA linear EST 01-MAY-2001
LOCUS 602638412F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:476234 5',
DEFINITION mRNA sequence.
ACCESSION BG686133
VERSION BG686133.1 GI:13917530
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 743)
NIH-MGC http://imgc.ncbi.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LLCM1626 row: g column: 03
High quality sequence stop: 740.
Location/Qualifiers

FEATURES

source

1..743

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4766234"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_48"
/notes="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
```

ORIGIN

```
Query Match      100.0%; Score 271; DB 12; Length 743;
Best Local Similarity 100.0%; Pred.No. 2.8e-47;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGCTGCACGATGATGCTCTCCCTACATCTCGAATGGACCTAGACCT 60
DB 234 AACGGCTGCACGATGATGCTCTCCCTACATCTCGAATGGACCTAGACCT 293
QY 61 GGGCGGTGTACCGCGCTCACTGTTCACTCTCGAGGCCCTGTACGACTGTGCCGA 120
DB 294 GGGCGGTGTACCGCGCTCACTGTTCACTCTCGAGGCCCTGTACGACTGTGCCGA 353
QY 121 CATGTGGCCGACCTTTCTGTGGAGGAAACCCCACTCAGTGTAGAGATCTTCAACCGCGGC 180
DB 354 CATGTGGCCGACCTTTCTGTGGAGGAAACCCCACTCAGTGTAGAGATCTTCAACCGCGGC 413
QY 181 CTCTACTTCTGTAGAGACCGCAAGGCTGAGCGGCGCTGCGCGGTGACCGCGGC 240
DB 414 CTCTACTTCTGTAGAGACCGCAAGGCTGAGCGGCGCTGCGCGGTGACCGCGGC 473
QY 241 GGGGTGCAATAGCCATCATGACCTTCAAG 271
DB 474 GGGGTGCAATAGCCATCATGACCTTCAAG 504
```

```
RESULT 4
EG757089      820 bp      mRNA      linear      EST 15-MAY-2001
LOCUS         602715124F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4855517 5',
DEFINITION    mRNA sequence.
ACCESSION     EG757089
VERSION       EG757089.1 GI:14067742
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE     1 (bases 1 to 820)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cga@bbs.fda.gov
              Tissue Procurement: Louis W. Staudt, M.D., Ph.D.
              cDNA Library Preparation: Ling Hong/Rubin Laboratory
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LMC1704 row: 0 column: 06
              High quality sequence stop: 675.
              Location/Qualifiers
                1..820
                /organism="Homo sapiens"
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4855517"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_48"
/notes="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
```

ORIGIN

```
Query Match      100.0%; Score 271; DB 12; Length 820;
Best Local Similarity 100.0%; Pred.No. 2.9e-47;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGCTGCACGATGATGCTCTCCCTACATCTCGAATGGACCTAGACCT 60
DB 241 AACGGCTGCACGATGATGCTCTCCCTACATCTCGAATGGACCTAGACCT 300
QY 61 GGGCGGTGTACCGCGCTCACTGTTCACTCTCGAGGCCCTGTACGACTGTGCCGA 120
DB 301 GGGCGGTGTACCGCGCTCACTGTTCACTCTCGAGGCCCTGTACGACTGTGCCGA 360
QY 121 CATGTGGCCGACCTTTCTGTGGAGGAAACCCCACTCAGTGTAGAGATCTTCAACCGCGGC 180
DB 361 CATGTGGCCGACCTTTCTGTGGAGGAAACCCCACTCAGTGTAGAGATCTTCAACCGCGGC 420
QY 181 CTCTACTTCTGTAGAGACCGCAAGGCTGAGCGGCGCTGCGCGGTGACCGCGGC 240
DB 421 CTCTACTTCTGTAGAGACCGCAAGGCTGAGCGGCGCTGCGCGGTGACCGCGGC 480
QY 241 GGGGTGCAATAGCCATCATGACCTTCAAG 271
DB 481 GGGGTGCAATAGCCATCATGACCTTCAAG 511
```

```
RESULT 5
BX464579      853 bp      mRNA      linear      EST 22-MAY-2003
LOCUS         BX464579 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
DEFINITION    Clone CS0D0003YB14 5-PRIME, mRNA sequence.
ACCESSION     BX464579
VERSION       BX464579.1 GI:31031641
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE     1 (bases 1 to 853)
AUTHORS      L.M.B., Gruber,C., Jesse,J. and Polyes,D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL       Unpublished (2001)
COMMENT       Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 101 91006 EVRY cedex - France
              Email: seq@genoscope.cns.fr, Web: www.genoscope.cns.fr
              Library was constructed by Life Technologies, a division of
              Invitrogen. This sequence belongs to sequence cluster 6672.r For
              more information about this cluster, see
              http://www.genoscope.cns.fr/
              cgi-bin/cluster.cgi?seq=CSID0001ZF10Q1&cluster=6672.r. Contact :
              Feng Liang Email: fliang@lifetech.com URL:
              http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
              Paradise Avenue Genoscope sequence ID: CSID0001ZF10Q1.
              Location/Qualifiers
                1..853
                /organism="Homo sapiens"
```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODG003YB14"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 100.0%; Score 271; DB 13; Length 853;
Best Local Similarity 100.0%; Pred. No. 2.9e-47;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGCTGCCACGTGAATTGCTCTTCTCCGCTACATCTCGACTGGAAGTGAACCT 60
DB 242 AACGGCTGCCACGTGAATTGCTCTTCTCCGCTACATCTCGACTGGAAGTGAACCT 301
QY 61 GGGCGCTGCTACCGGCTCACTGTTCACTCTTGGAGCCCTGCTACGACTGTGCCGA 120
DB 302 GGGCGCTGCTACCGGCTCACTGTTCACTCTTGGAGCCCTGCTACGACTGTGCCGA 361
QY 121 CATGTGGCCGACTTTCTGGAGAGGAACCCCAACCTCAGTCTGAGATCTTGACGGCGGC 180
DB 362 CATGTGGCCGACTTTCTGGAGAGGAACCCCAACCTCAGTCTGAGATCTTGACGGCGGC 421
QY 181 CTCTACTTCTGTGAGGACCGCAAGCTGAGCCGAGGGGCTGCGCGCTGCAACCGCGCC 240
DB 422 CTCTACTTCTGTGAGGACCGCAAGCTGAGCCGAGGGGCTGCGCGCTGCAACCGCGCC 481
QY 241 GGGGTGCAATAGCCATCATGACCTTCAAG 271
DB 482 GGGGTGCAATAGCCATCATGACCTTCAAG 512

RESULT 6 872 bp mRNA linear EST 15-MAY-2001
BG758510 60212721F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4851069 5',
LOCUS mRNA sequence.
DEFINITION
ACCESSION BG758510.1 GI:14069163
VERSION
KEYWORDS EST.
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS 1 (bases 1 to 872)
TITLE NIH-MGC
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LCM1698 row: 1 column: 06
High quality sequence stop: 836.
Location/Qualifiers

FEATURES

1..872
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4851069"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"

ORIGIN

Query Match 100.0%; Score 271; DB 12; Length 872;
Best Local Similarity 100.0%; Pred. No. 2.9e-47;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGCTGCCACGTGAATTGCTCTTCTCCGCTACATCTCGACTGGAAGTGAACCT 60
DB 222 AACGGCTGCCACGTGAATTGCTCTTCTCCGCTACATCTCGACTGGAAGTGAACCT 281
QY 61 GGGCGCTGCTACCGGCTCACTGTTCACTCTTGGAGCCCTGCTACGACTGTGCCGA 120
DB 282 GGGCGCTGCTACCGGCTCACTGTTCACTCTTGGAGCCCTGCTACGACTGTGCCGA 341
QY 121 CATGTGGCCGACTTTCTGGAGAGGAACCCCAACCTCAGTCTGAGATCTTGACGGCGGC 180
DB 342 CATGTGGCCGACTTTCTGGAGAGGAACCCCAACCTCAGTCTGAGATCTTGACGGCGGC 401
QY 181 CTCTACTTCTGTGAGGACCGCAAGCTGAGCCGAGGGGCTGCGCGCTGCAACCGCGCC 240
DB 402 CTCTACTTCTGTGAGGACCGCAAGCTGAGCCGAGGGGCTGCGCGCTGCAACCGCGCC 461
QY 241 GGGGTGCAATAGCCATCATGACCTTCAAG 271
DB 462 GGGGTGCAATAGCCATCATGACCTTCAAG 492

RESULT 7 889 bp mRNA linear EST 01-MAY-2001
BG686876 602650861F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4763247 5',
LOCUS mRNA sequence.
DEFINITION
ACCESSION BG686876
VERSION BG686876.1 GI:13918273
KEYWORDS EST.
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS 1 (bases 1 to 889)
TITLE NIH-MGC
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LCM1618 row: 1 column: 16
High quality sequence start: 6
High quality sequence stop: 727.
Location/Qualifiers

FEATURES

1..889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4763247"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"

/note="Organ: NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 271; DB 12; Length 889;
Best Local Similarity 100.0%; Pred. No. 2.9e-47;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGCTGCACGATGGAATTGCTCTTCCTCCGCTACATCTGGAGACTGGAGACTAGACCT 60
DB 194 AACGGCTGCACGATGGAATTGCTCTTCCTCCGCTACATCTGGAGACTGGAGACTAGACCT 253
QY 61 GGGCGCTGCTACCGCGCTACCTGTTCACTCTGGAGCCCTGCTACGACTGTGCCGA 120
DB 254 GGGCGCTGCTACCGCGCTACCTGTTCACTCTGGAGCCCTGCTACGACTGTGCCGA 313
QY 121 CATGTGGCCGACTTTCTGCGAGGGAACCCCACTCAGTCTGAGATCTTCAACCGCGGC 180
DB 314 CATGTGGCCGACTTTCTGCGAGGGAACCCCACTCAGTCTGAGATCTTCAACCGCGGC 373
QY 181 CTCTACTTCTGTAGAGACCGCAAGGCTGAGCCGAGGCGCTGCGCGCTGACCGCGCC 240
DB 374 CTCTACTTCTGTAGAGACCGCAAGGCTGAGCCGAGGCGCTGCGCGCTGACCGCGCC 433
QY 241 GGGGTGCAATAGCCATCATGACTTCAAG 271
DB 434 GGGGTGCAATAGCCATCATGACTTCAAG 464

RESULT 8
B0065440 953 bp mRNA linear EST 02-APR-2002
LOCUS B0065440
DEFINITION AGENCOURT_6855061 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5929977.
5', mRNA sequence.
ACCESSION B0065440
VERSION B0065440.1 GI:19894486
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 953)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2108 row: P column: 10
High quality sequence stop: 634.
Location/Qualifiers

FEATURES

source

1. 953
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5929977"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"

/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

ORIGIN

Query Match 100.0%; Score 271; DB 13; Length 953;
Best Local Similarity 100.0%; Pred. No. 3e-47;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGCTGCACGATGGAATTGCTCTTCCTCCGCTACATCTGGAGACTGGAGACTAGACCT 60
DB 220 AACGGCTGCACGATGGAATTGCTCTTCCTCCGCTACATCTGGAGACTGGAGACTAGACCT 279
QY 61 GGGCGCTGCTACCGCGCTACCTGTTCACTCTGGAGCCCTGCTACGACTGTGCCGA 120
DB 280 GGGCGCTGCTACCGCGCTACCTGTTCACTCTGGAGCCCTGCTACGACTGTGCCGA 339
QY 121 CATGTGGCCGACTTTCTGCGAGGGAACCCCACTCAGTCTGAGATCTTCAACCGCGGC 180
DB 340 CATGTGGCCGACTTTCTGCGAGGGAACCCCACTCAGTCTGAGATCTTCAACCGCGGC 399
QY 181 CTCTACTTCTGTAGAGACCGCAAGGCTGAGCCGAGGCGCTGCGCGCTGACCGCGCC 240
DB 400 CTCTACTTCTGTAGAGACCGCAAGGCTGAGCCGAGGCGCTGCGCGCTGACCGCGCC 459
QY 241 GGGGTGCAATAGCCATCATGACTTCAAG 271
DB 460 GGGGTGCAATAGCCATCATGACTTCAAG 490

RESULT 9
B0055935 1052 bp mRNA linear EST 29-MAR-2002
LOCUS B0055935
DEFINITION AGENCOURT_6796291 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5808181
5', mRNA sequence.
ACCESSION B0055935
VERSION B0055935.1 GI:19815262
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1052)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2051 row: M column: 14
High quality sequence stop: 665.
Location/Qualifiers

FEATURES

source

1. 1052
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5808181"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:"

ECOR1: cDNA made by oligo-dt priming. Directionally cloned into Ecor1/Xho1 sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 271; DB 13; Length 1052;
Best Local Similarity 100.0%; Pred. No. 3.1e-47;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGTGCACAGTGAATTGCTCTTCTCCGCTACATCTGAGACTGGAGCTAGACCTT 60
DB 220 AACGGTGCACAGTGAATTGCTCTTCTCCGCTACATCTGAGACTGGAGCTAGACCTT 279
QY 61 GGCCGCTGTACCGGCTACCTGTTCACTCTGAGAGCCCTGCTAGACTGTGCCCCGA 120
DB 280 GGCCGCTGTACCGGCTACCTGTTCACTCTGAGAGCCCTGCTAGACTGTGCCCCGA 339
QY 121 CATGTGCGCAGCTTTCTGTGAGAGGAACCCCACTCACTGAGAGATCTTCAACCGCGCC 180
DB 340 CATGTGCGCAGCTTTCTGTGAGAGGAACCCCACTCACTGAGAGATCTTCAACCGCGCC 339
QY 181 CTCTACTTCTGTGAGAGACCGCAAGGCTGAGCCGAGGGGCTGCGGCTGACCGCGCC 240
DB 400 CTCTACTTCTGTGAGAGACCGCAAGGCTGAGCCGAGGGGCTGCGGCTGACCGCGCC 459
QY 241 GGGGTGCAATAGCCATCATGACTTTCAAG 271
DB 460 GGGGTGCAATAGCCATCATGACTTTCAAG 490

RESULT 10
AL559877
LOCUS AL559877 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
DEFINITION clone CS0D003YB14 5-PRIME, mRNA sequence.
ACCESSION AL559877
VERSION AL559877.2 GI:31284408
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:12505793.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr; Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6672.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D003DA070P1&cluster=6672.r. Contact :
Feng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D003DA070P1.

FEATURES
source

1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D003YB14"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcorV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 100.0%; Score 271; DB 9; Length 1201;
Best Local Similarity 100.0%; Pred. No. 3.2e-47;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGTGCACAGTGAATTGCTCTTCTCCGCTACATCTGAGACTGGAGCTAGACCTT 60
DB 248 AACGGTGCACAGTGAATTGCTCTTCTCCGCTACATCTGAGACTGGAGCTAGACCTT 307
QY 61 GGCCGCTGTACCGGCTACCTGTTCACTCTGAGAGCCCTGCTAGACTGTGCCCCGA 120
DB 308 GGCCGCTGTACCGGCTACCTGTTCACTCTGAGAGCCCTGCTAGACTGTGCCCCGA 367
QY 121 CATGTGCGCAGCTTTCTGTGAGAGGAACCCCACTCACTGAGAGATCTTCAACCGCGCC 180
DB 368 CATGTGCGCAGCTTTCTGTGAGAGGAACCCCACTCACTGAGAGATCTTCAACCGCGCC 427
QY 181 CTCTACTTCTGTGAGAGACCGCAAGGCTGAGCCGAGGGGCTGCGGCTGACCGCGCC 240
DB 428 CTCTACTTCTGTGAGAGACCGCAAGGCTGAGCCGAGGGGCTGCGGCTGACCGCGCC 487
QY 241 GGGGTGCAATAGCCATCATGACTTTCAAG 271
DB 488 GGGGTGCAATAGCCATCATGACTTTCAAG 518

RESULT 11
BX402063
LOCUS BX402063 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0D012YD18 5-PRIME, mRNA sequence.
ACCESSION BX402063
VERSION BX402063.1 GI:30626645
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr; Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6672.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D012DB09P1&cluster=6672.r. Contact :
Feng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D012DB09P1.

FEATURES
source

1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D012YD18"
/cell_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcorV

ORIGIN

sites of the PCMVSPORT 6 vector. Library was normalized."

Query Match 100.0%; Score 271; DB 13; Length 1201;
Best Local Similarity 100.0%; Pred. No. 3.2e-47;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGGCTGCCACGTGGAATGCTCTTCCCTCGCTACATCTCGAGACTGGAACCTAGACCT 60
DB 308 AACGGCTGCCACGTGGAATGCTCTTCCCTCGCTACATCTCGAGACTGGAACCTAGACCT 367
QY 61 GGGCGCTGCTACCGCGCTCACTGCTTCACTCTGAGAGCCCTGCTACGACTGTGCCGA 120
DB 368 GGGCGCTGCTACCGCGCTCACTGCTTCACTCTGAGAGCCCTGCTACGACTGTGCCGA 427
QY 121 CATGTGGCCGACTTTCTGCGAGGGAACCCCACTGATGTGAGATCTTCCACCGCGCGC 180
DB 428 CATGTGGCCGACTTTCTGCGAGGGAACCCCACTGATGTGAGATCTTCCACCGCGCGC 487
QY 181 CTCTACTTCTGTGAGAGACCGCAAGCTGAGCCCGAGGGGCTGGCGGCTGCAACCGCGC 240
DB 488 CTCTACTTCTGTGAGAGACCGCAAGCTGAGCCCGAGGGGCTGGCGGCTGCAACCGCGC 547
QY 241 GGGGTGCAATATGACCATCATGACCTTCAAG 271
DB 548 GGGGTGCAATATGACCATCATGACCTTCAAG 578

RESULT 12

BF975166 942 bp mRNA linear EST 22-JAN-2001
LOCUS 602244657F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4335639 5',
DEFINITION mRNA sequence.

ACCSSION BF975166.1 GI:12342381
VERSION BF975166
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 942)
JOURNAL NIH-MGC http://imgc.ncbi.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM1207 row: a column: 16
High quality sequence stop: 707.
Location/Qualifiers

FEATURES

source

1. 942

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4335639"

/tissue_type="primary B-cells from tonsils (cell line)"

/lab_host="DH10B (phage-resistant)"

/clone_id="NIH MGC 48"

/note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 99.4%; Score 269.4; DB 10; Length 942;
Best Local Similarity 99.6%; Pred. No. 6.5e-47;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AACGGCTGCCACGTGGAATGCTCTTCCCTCGCTACATCTTGGAGCTGGAGACTTAGACCT 60
DB 235 AACGGCTGCCACGTGGAATGCTCTTCCCTCGCTACATCTTGGAGCTGGAGACTTAGACCT 294
QY 61 GGGCGCTGCTACCGCGCTCACTGCTTCACTCTGAGAGCCCTGCTACGACTGTGCCGA 120
DB 295 GGGCGCTGCTACCGCGCTCACTGCTTCACTCTGAGAGCCCTGCTACGACTGTGCCGA 354
QY 121 CATGTGGCCGACTTTCTGCGAGGGAACCCCACTGATGTGAGATCTTCCACCGCGCGC 180
DB 355 CATGTGGCCGACTTTCTGCGAGGGAACCCCACTGATGTGAGATCTTCCACCGCGCGC 414
QY 181 CTCTACTTCTGTGAGAGACCGCAAGCTGAGCCCGAGGGGCTGGCGGCTGCAACCGCGC 240
DB 415 CTCTACTTCTGTGAGAGACCGCAAGCTGAGCCCGAGGGGCTGGCGGCTGCAACCGCGC 474
QY 241 GGGGTGCAATATGACCATCATGACCTTCAAG 271
DB 475 GGGGTGCAATATGACCATCATGACCTTCAAG 505

RESULT 13

CD707143 535 bp mRNA linear EST 25-JUN-2003
LOCUS CD707143
DEFINITION EST23670 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCSSION CD707143 GI:32237773
VERSION CD707143.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 535)
JOURNAL Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and
Zeng, Y.-X.
COMMENT Transcriptional Gene Expression Profile of Human Nasopharynx
Unpublished (2003)
Contact: Yixin Zeng
Cancer Center
Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 51060, China
Tel: 86-1380-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@gzsun.edu.cn.
Location/Qualifiers

FEATURES

source

1. 535

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="normal nasopharynx"

/clone_id="human nasopharynx"

/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"

ORIGIN

Query Match 95.0%; Score 257.4; DB 14; Length 535;
Best Local Similarity 99.6%; Pred. No. 1.9e-44;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AACGGCTGCCACGTGGAATGCTCTTCCCTCGCTACATCTTGGAGCTGGAGACTTAGACCT 60
DB 277 AACGGCTGCCACGTGGAATGCTCTTCCCTCGCTACATCTTGGAGCTGGAGACTTAGACCT 336
QY 61 GGGCGCTGCTACCGCGCTCACTGCTTCACTCTGAGAGCCCTGCTACGACTGTGCCGA 120
DB 337 GGGCGCTGCTACCGCGCTCACTGCTTCACTCTGAGAGCCCTGCTACGACTGTGCCGA 396

QY 121 CATGTGCGCAGCTTTCTGCGAGGAAACCCCACTGATGTAAGATCTTACCGCGCCG 180
 DB 397 CGGTGCGCCAGCTTTCTGCGAGGAAACCCCACTGATGTAAGATCTTACCGCGCCG 456
 QY 181 CTCTACTCTGTGTAGGAGCCGCAAGCGTGAACCCCGAGGGGTGTGGGGCTGACCGCGCCG 240
 DB 457 CTCTACTCTGTGTAGGAGCCGCAAGCGTGAACCCCGAGGGGTGTGGGGCTGACCGCGCC 516
 QY 241 GGGGTGCAATATGACCATCA 259
 DB 517 GGGGTGCAATATGACCATCA 535

RESULT 14
 BGI44705 522 bp mRNA linear EST 01-FEB-2001
 LOCUS U73107.y1 Soares_mouse_MMBG_bcell Mus musculus cDNA clone
 DEFINITION IMAGE:333637 5' similar to TR:Q9WE0 Q9WE0 ACTIVATION-INDUCED
 CYTIDINE DEAMINASE, mRNA sequence.
 ACCESSION BGI44705 GI:12648105
 VERSION BGI44705
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 522)
 REFERENCE NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:1077801
 Seq primer: -40RP from Gibco
 High quality sequence stop: 487.
 Location/Qualifiers
 1..522
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:333637"
 /lab_host="DH10B (phage-resistant)"
 /clone_1ib="Soares_mouse_MMBG_bcell"
 /note="Organ: germinal B-cell; Vector: pT7T3D-Pac
 (Pharmacia) with a modified polylinker Site 1: Not I;
 Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer [5'
 TGTTACCAATCTGAAGTGAAGGAGCGCGCTGTGTTTTTTTTTTTTTTTTTTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pT7T3 vector.
 Library is normalized; constructed by Bento Soares and
 M.Fatima Bonaudo."

ORIGIN
 Query Match 81.8%; Score 221.6; DB 12; Length 522;
 Best Local Similarity 89.2%; Pred. No. 7e-37;
 Matches 239; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 4 GGGTGCACGTGGAATGCTCTTCCCGCTACATCTGGAAGTGGACCTAGACCTGAC 63
 DB 77 GGGTGCACGTGGAATGCTCTTCCACGCTACATCTGGAAGTGGACCTAGACCTGAC 136
 QY 64 CGCTGCTACCGCGCTACCTGCTTCACTCTGGAAGCCCTGCTACACACTGTGCGGACAT 123
 DB 137 CGGTGTATACCGCGCTACCTGCTTCACTCTGGAAGCCCGGCTATGACTGTGCGGAC 196
 QY 124 GTGGCGCACTTCTGAGAGGAAACCCCACTGAGCTGAGATCTTACCGGCGGCGCC 183
 DB 197 GTGGCGCACTTCTGAGAGTGAACCTTAACTCAAGCTGAGGATTTTACCGGCGGCGCC 256

QY 184 TACTTCTGTAGAGACCCGCAAGGCTGAGCCCGAGGGCTGCGGGCTGCAACCGCGCCG 243
 DB 257 TACTTCTGTAGAGACCCGCAAGGCTGAGCCCGAGGGCTGCGGGCTGCAACCGCGCGG 316
 QY 244 GTCCAGATCGGAGATCATGACCTTCAAG 271
 DB 317 GTCCAGATCGGAGATCATGACCTTCAAG 344

RESULT 15
 B637360 653 bp mRNA linear EST 26-OCT-2001
 LOCUS B637360
 DEFINITION B637360 RIKEN full-length enriched, adult male aorta and vein Mus
 musculus cDNA clone A530070C03 5', mRNA sequence.
 ACCESSION B637360 GI:16473145
 VERSION B637360.1
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 653)
 REFERENCE Aikawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
 Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
 Kono,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
 Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
 Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
 Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Aikawa,T., et al. 2001)
 JOURNAL Unpublished (2001)
 COMMENT Contact: Yoshinide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.go.jp,
 URL:<http://genome.gsc.riken.go.jp/>
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
 Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
 and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
 Sugahara,Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamahata,I.,
 Azawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
 Hayashizaki,Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 e mouse tissues.
 Location/Qualifiers
 1..653
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="A530070C03"
 /sex="male"

FEATURES
 source

```

/tissue_type="aorta and vein"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male aorta
and vein"

```

/note="Site 1: Sall; Site 2 BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genetic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGGAGAAGAAATCCCAACAGCTCTTTTTTTTTTTTTNN 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 455.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGGAGAAGATTCAGTAATAAATATATACCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pbluescript KS(+) after bulk excision from Lambda phage I."

ORIGIN

| | | | | |
|-----------------------|--------------|--------------------|----------------|-------------|
| Query Match | 81.8%; | Score 221.6; | DB 10; | Length 653; |
| Best Local Similarity | 89.2%; | Pred. No. 7.5e-37; | | |
| Matches 239; | Conservative | 0; | Mismatches 29; | Indels 0; |
| | | | Gaps | 0; |

[illegible]

Search completed: March 13, 2004, 00:39:03
Job time : 864.058 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 12:27:59 ; Search time 365.199 Seconds
(without alignments)
13767.261 Million cell updates/sec

Title: US-09-966-880A-14

Sequence: 1 attatttttactgtctgtaatt.....agcttcggcgatcctcttttg 116

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

GenBank: 1: gb_ba: 2: gb_hcg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_un: 14: gb_vl: 15: em_ba: 16: em_fun: 17: em_hum: 18: em_in: 19: em_mu: 20: em_om: 21: em_or: 22: em_ov: 23: em_pat: 24: em_ph: 25: em_pl: 26: em_ro: 27: em_sts: 28: em_un: 29: em_vl: 30: em_hcg_hum: 31: em_hcg_inv: 32: em_hcg_other: 33: em_hcg_mus: 34: em_hcg_pln: 35: em_hcg_rtd: 36: em_hcg_man: 37: em_hcg_vtl: 38: em_hcg_hum: 39: em_hcg_mus: 40: em_hcg_mus: 41: em_hcg_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | * Query Match | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------|
| 1 | 116 | 100.0 | 116 | 6 | BD016839 |
| 2 | 116 | 100.0 | 335 | 6 | AS577811 |
| 3 | 116 | 100.0 | 6564 | 6 | BD016835 |
| 4 | 116 | 100.0 | 11204 | 6 | BD016860 |
| 5 | 116 | 100.0 | 11204 | 6 | AB040430 |
| 6 | 116 | 100.0 | 71132 | 9 | AC092184 |
| 7 | 114.4 | 98.6 | 577 | 12 | AF529840 |
| 8 | 114.4 | 98.6 | 591 | 12 | AF529855 |
| 9 | 114.4 | 98.6 | 596 | 9 | AF529816 |
| 10 | 114.4 | 98.6 | 596 | 9 | AF529818 |
| 11 | 114.4 | 98.6 | 596 | 9 | AF529819 |
| 12 | 114.4 | 98.6 | 596 | 9 | AF529820 |
| 13 | 114.4 | 98.6 | 596 | 9 | AF529821 |
| 14 | 114.4 | 98.6 | 596 | 9 | AF529822 |
| 15 | 114.4 | 98.6 | 596 | 9 | AF529823 |
| 16 | 114.4 | 98.6 | 596 | 9 | AF529824 |
| 17 | 114.4 | 98.6 | 596 | 9 | AF529825 |
| 18 | 114.4 | 98.6 | 596 | 9 | AF529826 |
| 19 | 114.4 | 98.6 | 596 | 9 | AF529827 |
| 20 | 114.4 | 98.6 | 596 | 12 | AF529829 |
| 21 | 114.4 | 98.6 | 596 | 12 | AF529830 |
| 22 | 114.4 | 98.6 | 596 | 12 | AF529833 |
| 23 | 114.4 | 98.6 | 596 | 12 | AF529835 |
| 24 | 114.4 | 98.6 | 596 | 12 | AF529836 |
| 25 | 114.4 | 98.6 | 596 | 12 | AF529837 |
| 26 | 114.4 | 98.6 | 596 | 12 | AF529838 |
| 27 | 114.4 | 98.6 | 596 | 12 | AF529839 |
| 28 | 114.4 | 98.6 | 597 | 9 | BT007402 |
| 29 | 114.4 | 98.6 | 597 | 12 | AF529828 |
| 30 | 114.4 | 98.6 | 597 | 12 | AF529841 |
| 31 | 114.4 | 98.6 | 597 | 12 | AF529842 |
| 32 | 114.4 | 98.6 | 597 | 12 | AF529844 |
| 33 | 114.4 | 98.6 | 597 | 12 | AF529845 |
| 34 | 114.4 | 98.6 | 597 | 12 | AF529846 |
| 35 | 114.4 | 98.6 | 597 | 12 | AF529847 |
| 36 | 114.4 | 98.6 | 597 | 12 | AF529848 |
| 37 | 114.4 | 98.6 | 597 | 12 | AF529849 |
| 38 | 114.4 | 98.6 | 597 | 12 | AF529850 |
| 39 | 114.4 | 98.6 | 597 | 12 | AF529852 |
| 40 | 114.4 | 98.6 | 597 | 12 | AF529854 |
| 41 | 114.4 | 98.6 | 597 | 12 | BT008226 |
| 42 | 114.4 | 98.6 | 1828 | 9 | BC006296 |
| 43 | 114.4 | 98.6 | 2791 | 9 | AB040431 |
| 44 | 114.4 | 98.6 | 2818 | 6 | BD016833 |
| 45 | 112.8 | 97.2 | 547 | 12 | AF529856 |

ALIGNMENTS

RESULT 1
BD016839
LOCUS: BD016839 Novel cytidine deaminase. 116 bp DNA linear PAT 27-AUG-2002
DEFINITION: BD016839
ACCESSION: BD016839.1 GI:22556015
VERSION: UP 2001245669-A/12.
KEYWORDS: UP 2001245669-A/12.
SOURCE: Homo sapiens (human)
ORGANISM: Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE: Honjo, T. and Muramatsu, M.
TITLE: Novel cytidine deaminase
JOURNAL: Patent: UP 2001245669-A 12 11-SEP-2001;

COMMENT JAPAN TOBACCO INC. TASUKU HONJO
OS Homo sapiens (human)
PN JP 2001245669-A/12
PD 11-SEP-2001
PF 28-MAR-2000 JP 2000092981
PI TASUKU HONJO, MASAMICHI MURAMATSU
PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12,
PC A61P17/00,
PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12N9/78, C12P21/02, C12P21/08, C12N1/21, C12R1/19, PC
(C12N5/10, C12R1/91), C12N15/00, C12N5/00, C12N5/00, C12R1/91) CC
FH Key Location/Qualifiers
FEATURES
source 1..116
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 116; DB 6; Length 116;
Best Local Similarity 100.0%; Pred. No. 2.6e-27;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATATTTTCTGCTGGAATCTTTGTAGAAAACGAGAACTTTCAAGCCTGGG 60
1 ATATTTTCTGCTGGAATCTTTGTAGAAAACGAGAACTTTCAAGCCTGGG 60
DB 1 ATATTTTCTGCTGGAATCTTTGTAGAAAACGAGAACTTTCAAGCCTGGG 60
QY 61 AAGGCTGCATGAAATTCAGTTCGCTCCAGACAGCTTCGGCGCATCTTTTG 116
61 AAGGCTGCATGAAATTCAGTTCGCTCCAGACAGCTTCGGCGCATCTTTTG 116
DB 61 AAGGCTGCATGAAATTCAGTTCGCTCCAGACAGCTTCGGCGCATCTTTTG 116
RESULT 2 335 bp mRNA linear pr1 22-JUL-2003
HSAS77811
LOCUS Homo sapiens partial mRNA for activation-induced cytidine deaminase
DEFINITION (Aid gene).
ACCESSION AU577811.1 GI:33145978
VERSION A577811.1
KEYWORDS activation-induced cytidine deaminase; Aid gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Takahar, P., Coker, H.A., Fear, D., Smurthwaite, L., Durham, S.R. and
Gould, H.O.,
Allelexen and activation-induced deaminase drive heavy-chain class
switching to IgH in the nasal mucosa of hay fever patients
Unpublished
2 (bases 1 to 335)
JOURNAL Takhar, P.
REFERENCE Direct Submission
AUTHORS Submitted (21-JUL-2003) Takhar P., Biomedical Sciences Randall
TITLE Centre, King's College London, New Hunt House, Guy's Campus, St
JOURNAL Thomas St, London, SE1 1UL, UNITED KINGDOM
FEATURES
source 1..335
/organism="Homo sapiens"
/mol_type="mRNA"
/isolation_source="nasal biopsy from inferior turbinate of
an allergic rhinitis patient"
/db_xref="taxon:9606"
/clone="NR02"
/sex="Female"
/cell_type="B lymphocyte"
/tissue_type="nasal mucosa"
/dev_stage="adult"
/country="United Kingdom"
1..335
/gene="Aid"
<1..>335
CDS
gene
CDS

/gene="Aid"
/function="editing (involved in somatic hypermutation and
class switch recombination)"
/codon_start=3
/evidence=experimental
/product="activation-induced cytidine deaminase"
/protein_id="CA512261.1"
/db_xref="GI:33145979"
/translation="DPRGRCYRVWPTMSRCPYCAHVDPLRGNNLSLRIFETRL
FEDRKAPPEGLRLRLRAVQVIAIMFKDIFYQNTFVNHETFRAMGLHNSVRL
SRQLRLIL"
ORIGIN
Query Match 100.0%; Score 116; DB 9; Length 335;
Best Local Similarity 100.0%; Pred. No. 2.7e-27;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATATTTTCTGCTGGAATCTTTGTAGAAAACGAGAACTTTCAAGCCTGGG 60
1 ATATTTTCTGCTGGAATCTTTGTAGAAAACGAGAACTTTCAAGCCTGGG 60
DB 220 ATATTTTCTGCTGGAATCTTTGTAGAAAACGAGAACTTTCAAGCCTGGG 279
QY 61 AAGGCTGCATGAAATTCAGTTCGCTCCAGACAGCTTCGGCGCATCTTTTG 116
61 AAGGCTGCATGAAATTCAGTTCGCTCCAGACAGCTTCGGCGCATCTTTTG 116
DB 280 AAGGCTGCATGAAATTCAGTTCGCTCCAGACAGCTTCGGCGCATCTTTTG 335
RESULT 3 6564 bp DNA linear PAT 27-AUG-2002
BD016835
LOCUS Novel cytidine deaminase.
DEFINITION BD016835
ACCESSION BD016835.1 GI:22558011
VERSION JP 2001245669-A/8.
KEYWORDS Novel cytidine deaminase
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 6564)
AUTHORS Honjo, T. and Muramatsu, M.
TITLE Novel cytidine deaminase
JOURNAL Patent: JP 2001245669-A 8 11-SEP-2001;
JOURNAL JAPAN TOBACCO INC. TASUKU HONJO
COMMENT OS Homo sapiens (human)
PN JP 2001245669-A/8
PD 11-SEP-2001
PF 28-MAR-2000 JP 2000092981
PI TASUKU HONJO, MASAMICHI MURAMATSU
PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12,
PC A61P17/00,
PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12N9/78, C12P21/02, C12P21/08, C12N1/21, C12R1/19, PC
(C12N5/10, C12R1/91), C12N15/00, C12N5/00, C12N5/00, C12R1/91) CC
FH Key Location/Qualifiers
FEATURES
source 1..6564
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 116; DB 6; Length 6564;
Best Local Similarity 100.0%; Pred. No. 3e-27;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATATTTTCTGCTGGAATCTTTGTAGAAAACGAGAACTTTCAAGCCTGGG 60
1 ATATTTTCTGCTGGAATCTTTGTAGAAAACGAGAACTTTCAAGCCTGGG 60
DB 3156 ATATTTTCTGCTGGAATCTTTGTAGAAAACGAGAACTTTCAAGCCTGGG 3215
QY 61 AAGGCTGCATGAAATTCAGTTCGCTCCAGACAGCTTCGGCGCATCTTTTG 116
61 AAGGCTGCATGAAATTCAGTTCGCTCCAGACAGCTTCGGCGCATCTTTTG 116
DB 3216 AAGGCTGCATGAAATTCAGTTCGCTCCAGACAGCTTCGGCGCATCTTTTG 3271

```

RESULT 4
BD016860
LOCUS      BD016860
DEFINITION Novel cytidine deaminase.
ACCESSION BD016860.1
VERSION    BD016860.1 GI:22558036
KEYWORDS   JP 2001245669-A/33.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  Eukaryote, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Euteleostomi; Primates; Catarrhini; Homnidae; Homo.
TITLE       Honjo, T. and Muramatsu, M.
JOURNAL     Patent: JP 2001245669-A 33 11-SEP-2001;
            JAPAN TOBACCO INC, TASUKU HONJO
COMMENT     OS Homo sapiens (human)
            PN JP 2001245669-A/33
            PD 11-SEP-2001
            PF 28-MAR-2000 JP 2000092981
            PI TASUKU HONJO, MASAMICHI MURAMATSU
            PC C12N15/09, A61K39/395, A61P1/00, A61P11/06, A61P13/12,
            PC A61P17/00,
            PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC
            C12N1/21,
            PC C12N5/10, C12N9/78, C12P21/02, C12P21/08, C12N1/21, C12R1/19, PC
            (C12N5/10, C12R1/91), C12N15/00, C12N5/00, C12N5/00, C12R1/91) CC
            FH Key Location/Qualifiers
FEATURES
source     1..11204
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 116; DB 6; Length 11204;
Best Local Similarity 100.0%; Pred. No. 3.1e-27;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 ATTATTTTACTGCTGGAAATCTTTGTAGAAAACGAGAAAGACTTCAAGCTGGG 60
DB      8371 ATTATTTTACTGCTGGAAATCTTTGTAGAAAACGAGAAAGACTTCAAGCTGGG 8430
QY      61 AAGGCTGCATGAATAATTCAGTCTCTCCAGACAGCTTCGGCGATCCTTTG 116
DB      8431 AAGGCTGCATGAATAATTCAGTCTCTCCAGACAGCTTCGGCGATCCTTTG 8486
RESULT 5
AB040430
LOCUS      AB040430
DEFINITION Homo sapiens A1D gene for activation-induced cytidine deaminase,
ACCESSION AB040430.1
VERSION    AB040430.1 GI:9988407
KEYWORDS   A1D; activation-induced cytidine deaminase.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Euteleostomi; Primates; Catarrhini; Homnidae; Homo.
TITLE       Muto, T., Muramatsu, M., Taniwaki, M., Kinoshita, K. and Honjo, T.
JOURNAL     Isolation, tissue distribution, and chromosomal localization of the
            human activation-induced cytidine deaminase (A1D) gene
COMMENT     Genomics 68 (1), 85-88 (2000)
MEDLINE    20408890
PUBMED     10950930
REFERENCE  Revy, P., Muto, T., Levy, Y., Geisemann, F., Plebani, A., Sanai, O.,
AUTHORS     Catalan, N., Forveille, M., Dufourcq-Lagelouse, R., Gennery, A.,
            Tzscan, I., Ersoy, F., Kayserili, H., Ugazio, A.G., Brousse, N.,
            Muramatsu, M., Notarangelo, L.D., Kinoshita, K., Honjo, T., Fischer, A.

```

```

TITLE       and Durandy, A.
JOURNAL     Activation-induced cytidine deaminase (A1D) deficiency causes the
MEDLINE    autosomal recessive form of the Hyper-IgM syndrome (HIGM2)
PUBMED     Cell 102 (5), 565-575 (2000)
REFERENCE  20460541
AUTHORS     Muto, T., Muramatsu, M., Taniwaki, M., Kinoshita, K. and Honjo, T.
TITLE       Direct Submission
JOURNAL     Submitted (18-MAR-2000) Tasuku Honjo, Kyoto University, Department
            of Medical Chemistry, Faculty of Medicine, Yoshida, Sakyo-ku,
            Kyoto, Kyoto 606-8501, Japan (E-mail: honjo@med.kyoto-u.ac.jp,
            Tel: 81-75-753-4371 (ex. 4371), Fax: 81-75-753-4388)
FEATURES
source     1..11204
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /join(521..528,6280..6427,7807..8077,8371..8486,8956..9009)
            /gene="A1D"
            /protein_id="BAB12720.1"
            /cds(521..528,6280..6427,7807..8077,8371..8486,8956..9009)
            /gene="A1D"
            /codon_start=1
            /product="activation-induced cytidine deaminase"
            /protein_id="BAB12720.1"
            /db_xref="GI:9988408"
            /translation="MDSLNNRRKELVQFNVRWAKRRETVLCYVKKRDSATPSL
            DFGILRNKGCHVELFLRITSDMDLPGRCYRTWTWTSPTCDKARHVDPLRGNP
            NLSRIFATRIYFCDRKADEGRIRHRRGVQALMTFDFYFCWTFVNEHRTK
            AMSGILHNSVRLSRQLRRILPLVEYVDLDAFRTGL"
ORIGIN
Query Match      100.0%; Score 116; DB 6; Length 11204;
Best Local Similarity 100.0%; Pred. No. 3.1e-27;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 ATTATTTTACTGCTGGAAATCTTTGTAGAAAACGAGAAAGACTTCAAGCTGGG 60
DB      8371 ATTATTTTACTGCTGGAAATCTTTGTAGAAAACGAGAAAGACTTCAAGCTGGG 8430
QY      61 AAGGCTGCATGAATAATTCAGTCTCTCCAGACAGCTTCGGCGATCCTTTG 116
DB      8431 AAGGCTGCATGAATAATTCAGTCTCTCCAGACAGCTTCGGCGATCCTTTG 8486
RESULT 6
AC092184
LOCUS      AC092184
DEFINITION BAC library complete sequence.
ACCESSION AC092184
VERSION    AC092184.7
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Euteleostomi; Primates; Catarrhini; Homnidae; Homo.
TITLE       Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
            Albrooks, S.L., Amarantunga, H.C., Are, J.R., Ayale, M., Banks, T.,
            Barbata, J., Benton, J., Birge, K., Blankenburg, K., Bonin, D.,
            Bouck, J., Bowie, S., Blevins, M., Brown, E., Brown, M., Bryant, N.P.,
            Bunay, K., Butcher, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
            Cartron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
            Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhury, I., Christopoulos, C.,
            Cleveland, C.D., Cox, C., Coyle, M.D., Dahorne, S.R., David, R.,
            Davila, M.L., Davis, C., Davy-Carroil, L., Dederich, D.A.,
            Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
            Douthett, K., Edgar, D., Edgar, H., Dugan-Rocha, S., Durbin, K.J.,
            Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emerling, S.,
            Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P.,
            Franz, P., Gabist, A., Gao, J., Garcia, A., Garner, T., Garza, N.,
            Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
            Hall, R.,

```

Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlik, P., Haves, A., Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B., Homel, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolyvet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kurehl, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H., Lozano, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Maronde, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Metzger, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mochabac, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L., Qules, M., Ren, Y., Rivers, M., Rojase, A., Rojudoan, I., Rolfe, M., Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shoshitari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabori, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansay, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umami, K., Vaequez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Kucherlapati, R., Weinstein, G., and Gibbs, R.

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 71132)
AUTHORS Worley, K.C.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 71132)
AUTHORS Worley, K.C.
TITLE Direct Submission
JOURNAL Submitted (18-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 4 (bases 1 to 71132)
AUTHORS Worley, K.C.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 5 (bases 1 to 71132)
AUTHORS Worley, K.C.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On May 25, 2002 this sequence version replaced gi:20901754. INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons

flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES

| source | QUALSTAT-REPORT |
|--|---------------------|
| 1..71132 | Location/Qualifiers |
| /organism="Homo sapiens" | |
| /mol_type="genomic DNA" | |
| /db_xref="taxon:9606" | |
| /chromosome="12" | |
| /clone="RP11-438L7" | |
| /complement(1..1448) | |
| /note="overlaps bases 1..1448 of clone AC092490" | |
| /function="clone overlap" | |
| 187..408 | |
| /standard_name="57233" | |
| 439..560 | |
| /standard_name="92005" | |
| 774..881 | |
| /rpt_family="MIR" | |
| 903..1190 | |
| /rpt_family="AluSg" | |
| 1191..1213 | |
| /rpt_family="AT_rich" | |
| 1391..1807 | |
| /standard_name="6612" | |
| 1744..1819 | |
| /standard_name="8198" | |
| 1966..2254 | |
| /rpt_family="AluSg" | |
| 3296..3328 | |
| /rpt_family="TTTC)n" | |
| 4097..4249 | |
| /rpt_family="AluSg" | |
| 4360..4654 | |
| /rpt_family="AluY" | |
| complement(5140..5262) | |
| /rpt_family="FLAM_C" | |
| 5275..5304 | |
| /rpt_family="AT_rich" | |
| 5305..5563 | |
| /rpt_family="AluO" | |
| 5564..5599 | |
| /rpt_family="AT_rich" | |
| 5644..5733 | |
| /rpt_family="MSTD" | |
| 5734..6012 | |
| /rpt_family="AluSg" | |
| 6020..6336 | |
| /rpt_family="AluSx" | |
| 6337..6629 | |
| /rpt_family="MSTD" | |
| 6630..6650 | |
| /rpt_family="CAA)n" | |
| complement(6711..6817) | |
| /rpt_family="MIR" | |
| complement(7033..7327) | |
| /rpt_family="AluSx" | |

```
source      1. .577
            /organism="Homo sapiens"
```

```

source
1..591
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
<1..591
/gene="AID"
<1..591
/cds
<1..591
/gene="AID"
/gene="AID"
/codon_start=1
/product="AID"
/protein_id="AA05442.1"
/db_xref="GI:22257298"

```

ORIGIN

Query Match 98.6%; Score 114.4; DB 12; Length 591;
Best Local Similarity 99.1%; Pred. No. 9.3e-27;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATATTCTTCTGCTGGAATCTTTGTAGAAAAACCAAGAACTTTCAAGCCTGGG 60
422 ATATTCTTCTGCTGGAATCTTTGTAGAAAAACCAAGAACTTTCAAGCCTGGG 481

Db 61 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCTTTTG 116
482 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCTTTTG 537

RESULT 9
AF529816 596 bp mRNA linear PRI 17-SEP-2002
LOCUS Homo sapiens clone Ramos 2 AID (AID) mRNA, partial cds.
AF529816 GI:22297219
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 596)
Somatic hypermutation of the AID transgene in B and non-B cells
Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
AUTHORS Martin, A. and Scharff, M.D.
DIRECT SUBMISSION
TITLE Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA
LOCATION/Qualifiers

FEATURES
source 1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 2"
1..596
/gene="AID"
1..596
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="P46954.3.1"
/db_xref="GI:22297220"
/translation="MDSLMNRKFLYQFNWAKGRRETYLCYVVKRDSATSPSL
DGYLRNKGCHVELFLRYISDMDDPBGCVATWPTGSRPCYDCARVADLRGNP
NLSIRITRALYFCEDEKAPBEGRLRHAGVOIAIMTFKDYFCWNTFVENHRTFK
AMEGLHNSVRLSRQLRIILPLRYEVDLRLDARITGL"

ORIGIN

Query Match 98.6%; Score 114.4; DB 9; Length 596;
Best Local Similarity 99.1%; Pred. No. 9.3e-27;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATATTCTTCTGCTGGAATCTTTGTAGAAAAACCAAGAACTTTCAAGCCTGGG 60
428 ATATTCTTCTGCTGGAATCTTTGTAGAAAAACCAAGAACTTTCAAGCCTGGG 487

Db 61 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCTTTTG 116
488 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCTTTTG 543

RESULT 11
AF529819 596 bp mRNA linear PRI 17-SEP-2002
LOCUS Homo sapiens clone Ramos 5 AID (AID) mRNA, partial cds.
AF529819 GI:22297225
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 596)
Somatic hypermutation of the AID transgene in B and non-B cells
Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
AUTHORS Martin, A. and Scharff, M.D.
DIRECT SUBMISSION
TITLE Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of

ORIGIN

Query Match 98.6%; Score 114.4; DB 9; Length 596;
Best Local Similarity 99.1%; Pred. No. 9.3e-27;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATATTCTTCTGCTGGAATCTTTGTAGAAAAACCAAGAACTTTCAAGCCTGGG 60
428 ATATTCTTCTGCTGGAATCTTTGTAGAAAAACCAAGAACTTTCAAGCCTGGG 487

Db 61 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCTTTTG 116
488 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCTTTTG 543

RESULT 11
AF529819 596 bp mRNA linear PRI 17-SEP-2002
LOCUS Homo sapiens clone Ramos 5 AID (AID) mRNA, partial cds.
AF529819 GI:22297225
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 596)
Somatic hypermutation of the AID transgene in B and non-B cells
Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
AUTHORS Martin, A. and Scharff, M.D.
DIRECT SUBMISSION
TITLE Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of

RESULT 10
AF529818 596 bp mRNA linear PRI 17-SEP-2002
LOCUS Homo sapiens clone Ramos 4 truncated AID (AID) mRNA, complete cds.
AF529818 GI:22297223
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 596)
Somatic hypermutation of the AID transgene in B and non-B cells
Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
AUTHORS Martin, A. and Scharff, M.D.
DIRECT SUBMISSION
TITLE Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA
LOCATION/Qualifiers

FEATURES
source 1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 4"
1..596
/gene="AID"
1..570
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="truncated AID"
/protein_id="P46954.3.1"
/db_xref="GI:22297224"
/translation="MDSLMNRKFLYQFNWAKGRRETYLCYVVKRDSATSPSL
DGYLRNKGCHVELFLRYISDMDDPBGCVATWPTGSRPCYDCARVADLRGNP
NLSIRITRALYFCEDEKAPBEGRLRHAGVOIAIMTFKDYFCWNTFVENHRTFK
AMEGLHNSVRLSRQLRIILPLRYEVDL"

ORIGIN

Query Match 98.6%; Score 114.4; DB 9; Length 596;
Best Local Similarity 99.1%; Pred. No. 9.3e-27;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATATTCTTCTGCTGGAATCTTTGTAGAAAAACCAAGAACTTTCAAGCCTGGG 60
428 ATATTCTTCTGCTGGAATCTTTGTAGAAAAACCAAGAACTTTCAAGCCTGGG 487

Db 61 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCTTTTG 116
488 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCTTTTG 543

RESULT 11
AF529819 596 bp mRNA linear PRI 17-SEP-2002
LOCUS Homo sapiens clone Ramos 5 AID (AID) mRNA, partial cds.
AF529819 GI:22297225
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 596)
Somatic hypermutation of the AID transgene in B and non-B cells
Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
AUTHORS Martin, A. and Scharff, M.D.
DIRECT SUBMISSION
TITLE Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of

ORIGIN

Query Match 98.6%; Score 114.4; DB 9; Length 596;
Best Local Similarity 99.1%; Pred. No. 9.3e-27;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATATTCTTCTGCTGGAATCTTTGTAGAAAAACCAAGAACTTTCAAGCCTGGG 60
428 ATATTCTTCTGCTGGAATCTTTGTAGAAAAACCAAGAACTTTCAAGCCTGGG 487

Db 61 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCTTTTG 116
488 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCTTTTG 543

RESULT 11
AF529819 596 bp mRNA linear PRI 17-SEP-2002
LOCUS Homo sapiens clone Ramos 5 AID (AID) mRNA, partial cds.
AF529819 GI:22297225
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 596)
Somatic hypermutation of the AID transgene in B and non-B cells
Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
AUTHORS Martin, A. and Scharff, M.D.
DIRECT SUBMISSION
TITLE Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of

FEATURES
source

Medicine, 1300 Morris Park Ave. Chantin 404, Bronx, NY 10461, USA

gene
CDS

1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 5"
1..>596
/gene="AID"
1..>596
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="AA05406.1"
/db_xref="GI:22297226"
/translation="MDSLNNRRKFLYQFNWAKGRRETYLCYVVKRDSATFSL
DGYLRNKGCHVELFLRYISDMDPGRCYRTWTWSPCYDCARHVADELKGNP
NLSIRIFPARLYFCVDRKAEPEGRLRHRAGVOIAIMTFKDYFCWNTFVENHRTFK
AMEGHNVSRLSRQLRILLPLYVDLDRARFTLGL"

ORIGIN

Query Match 98.6%; Score 114.4; DB 9; Length 596;
Best Local Similarity 99.1%; Pred. No. 9.3e-27;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAAATCTTTTGAGAAAACCAAGAACTTTCAAGCTGGG 60
DB 428 ATTATTTTACTGCTGGAAATCTTTTGAGAAAACCAAGAACTTTCAAGCTGGG 487
QY 61 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 116
DB 488 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 543

RESULT 12
LOCUS

AF529820 596 bp mRNA linear PRI 17-SEP-2002
DEFINITION Homo sapiens clone Ramos 6 AID (AID) mRNA, partial cds.
ACCESSION AF529820
VERSION AF529820.1 GI:22297227

KEYWORDS
SOURCE

Homo sapiens (human)
Homo sapiens

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Martin, A. and Schaff, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
2 (bases 1 to 596)
AUTHORS Martin, A. and Schaff, M.D.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Chantin 404, Bronx, NY 10461, USA

FEATURES
source

1..596
Location/Qualifiers

1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 6"
1..>596
/gene="AID"
1..>596
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="AA05407.1"
/db_xref="GI:22297228"
/translation="MDSLNNRRKFLYQFNWAKGRRETYLCYVVKRDSATFSL
DGYLRNKGCHVELFLRYISDMDPGRCYRTWTWSPCYDCARHVADELKGNP
NLSIRIFPARLYFCVDRKAEPEGRLRHRAGVOIAIMTFKDYFCWNTFVENHRTFK
AMEGHNVSRLSRQLRILLPLYVDLDRARFTLGL"

ORIGIN

AMEGHNVSRLSRQLRILLPLYVDLDRARFTLGL"

Query Match 98.6%; Score 114.4; DB 9; Length 596;
Best Local Similarity 99.1%; Pred. No. 9.3e-27;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAAATCTTTTGAGAAAACCAAGAACTTTCAAGCTGGG 60
DB 428 ATTATTTTACTGCTGGAAATCTTTTGAGAAAACCAAGAACTTTCAAGCTGGG 487
QY 61 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 116
DB 488 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 543

RESULT 13
LOCUS

AF529821 596 bp mRNA linear PRI 17-SEP-2002
DEFINITION Homo sapiens clone Ramos 7 AID (AID) mRNA, partial cds.
ACCESSION AF529821
VERSION AF529821.1 GI:22297229

KEYWORDS
SOURCE

Homo sapiens (human)
Homo sapiens

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Martin, A. and Schaff, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
2 (bases 1 to 596)
AUTHORS Martin, A. and Schaff, M.D.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Chantin 404, Bronx, NY 10461, USA

FEATURES
source

1..596
Location/Qualifiers

1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 7"
1..>596
/gene="AID"
1..>596
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="AA05408.1"
/db_xref="GI:22297230"
/translation="MDSLNNRRKFLYQFNWAKGRRETYLCYVVKRDSATFSL
DGYLRNKGCHVELFLRYISDMDPGRCYRTWTWSPCYDCARHVADELKGNP
NLSIRIFPARLYFCVDRKAEPEGRLRHRAGVOIAIMTFKDYFCWNTFVENHRTFK
AMEGHNVSRLSRQLRILLPLYVDLDRARFTLGL"

ORIGIN

1..596
Location/Qualifiers

1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 6"
1..>596
/gene="AID"
1..>596
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="AA05407.1"
/db_xref="GI:22297228"
/translation="MDSLNNRRKFLYQFNWAKGRRETYLCYVVKRDSATFSL
DGYLRNKGCHVELFLRYISDMDPGRCYRTWTWSPCYDCARHVADELKGNP
NLSIRIFPARLYFCVDRKAEPEGRLRHRAGVOIAIMTFKDYFCWNTFVENHRTFK
AMEGHNVSRLSRQLRILLPLYVDLDRARFTLGL"

QY 1 ATTATTTTACTGCTGGAAATCTTTTGAGAAAACCAAGAACTTTCAAGCTGGG 60
DB 428 ATTATTTTACTGCTGGAAATCTTTTGAGAAAACCAAGAACTTTCAAGCTGGG 487
QY 61 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 116
DB 488 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 543

RESULT 14
LOCUS

AF529822 596 bp mRNA linear PRI 17-SEP-2002
DEFINITION Homo sapiens clone Ramos 6 AID (AID) mRNA, partial cds.
ACCESSION AF529822
VERSION AF529822.1 GI:22297227

DEFINITION Homo sapiens clone Ramos 8 AID (AID) mRNA, partial cds.
ACCESSION AF529822
VERSION AF529822.1 GI:22297231
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Martin, A. and Scharff, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
REFERENCE
AUTHORS Martin, A. and Scharff, M.D.
TITLE Direct Submision
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Bronx, NY 10461, USA
FEATURES
LOCATION/Qualifiers
SOURCE
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 8"
1..>596
/gene="AID"
1..>596
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="AA095409.1"
/db_xref="GI:22297232"
/translation="MDSILNRRKFLYQFNVMWAKGRRETYLCYVVKRDSATSFSL
DFGYLRNKGCHVELFLRYISDMDDPGRCYVWTPTSPCYDCARHVAADPLRGNP
NLSLRIFTARLYPCEDRKAPESGLRLHRAVGQVIAIMTPDYFCNTPFVNHERRTFK
AMEGLHNSVRLSRQLRRIILPLYEVDDLDARFTLGL"
ORIGIN
Query Match 98.6%; Score 114.4; DB 9; Length 596;
Best Local Similarity 99.1%; Pred. No. 9.3e-27;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATTATTTTACTGCTGGAATACCTTTGTAGAAAAACGAAAGACTTTCAAAAGCTGGG 60
DB 428 ATTATTTTACTGCTGGAATACCTTTGTAGAAAAACGAAAGACTTTCAAAAGCTGGG 487
QY 61 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCCTTTTG 116
DB 488 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCCTTTTG 543
RESULT 15
AF529823 596 bp mRNA linear PRI 17-SEP-2002
DEFINITION Homo sapiens clone Ramos 9 AID (AID) mRNA, partial cds.
ACCESSION AF529823
VERSION AF529823.1 GI:22297233
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Martin, A. and Scharff, M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
REFERENCE
AUTHORS Martin, A. and Scharff, M.D.
TITLE Direct Submision
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Bronx, NY 10461, USA
FEATURES
LOCATION/Qualifiers
SOURCE
1..596

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 9"
1..>596
/gene="AID"
1..>596
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="AA095410.1"
/db_xref="GI:22297234"
/translation="MDSILNRRKFLYQFNVMWAKGRRETYLCYVVKRDSATSFSL
DFGYLRNKGCHVELFLRYISDMDDPGRCYVWTPTSPCYDCARHVAADPLRGNP
NLSLRIFTARLYPCEDRKAPESGLRLHRAVGQVIAIMTPDYFCNTPFVNHERRTFK
AMEGLHNSVRLSRQLRRIILPLYEVDDLDARFTLGL"
ORIGIN

Query Match 98.6%; Score 114.4; DB 9; Length 596;
Best Local Similarity 99.1%; Pred. No. 9.3e-27;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATTATTTTACTGCTGGAATACCTTTGTAGAAAAACGAAAGACTTTCAAAAGCTGGG 60
DB 428 ATTATTTTACTGCTGGAATACCTTTGTAGAAAAACGAAAGACTTTCAAAAGCTGGG 487
QY 61 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCCTTTTG 116
DB 488 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCCTTTTG 543

Search completed: March 12, 2004, 21:22:31
Cpu time : 365.199 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 12:26:39 ; Search time 46.4502 Seconds
(without alignments)
10609.014 Million cell updates/sec

Title: US-09-966-880A-14

Perfect score: 116
Sequence: 1 attactttcttcgtcgtgaat.....agcttcgcgcacatcttttg 116

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_29cand4:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 116 | 100.0 | 116 | 3 AAC5318 | AAC5318 Human act |
| 2 | 116 | 100.0 | 6564 | 3 AAC5314 | AAC5314 Human act |
| 3 | 116 | 100.0 | 11204 | 3 AAC5339 | AAC5339 Human act |
| 4 | 116 | 100.0 | 11204 | 6 AB873286 | AB873286 DNA encod |
| 5 | 114.4 | 98.6 | 597 | 8 ADB88952 | ADB88952 A1D gene |
| 6 | 114.4 | 98.6 | 1543 | 7 ABX05468 | ABX05468 Human nov |
| 7 | 114.4 | 98.6 | 2791 | 6 ABS73287 | ABS73287 DNA encod |
| 8 | 114.4 | 98.6 | 2791 | 6 ABS73288 | ABS73288 DNA encod |
| 9 | 114.4 | 98.6 | 2818 | 3 AAC5312 | AAC5312 Human act |
| 10 | 96.8 | 83.4 | 2440 | 3 AAC5307 | AAC5307 Mouse act |
| 11 | 34.2 | 29.5 | 1348 | 6 ABN96785 | ABN96785 Gene #328 |
| 12 | 34.2 | 29.5 | 1446 | 6 AAA12409 | AAA12409 CDNA encd |
| 13 | 34.2 | 29.5 | 1534 | 2 AA220856 | AA220856 Polynucle |
| 14 | 34.2 | 29.5 | 1534 | 4 AAS59293 | AAS59293 Human CDN |
| 15 | 34.2 | 29.5 | 1534 | 6 ABA90962 | ABA90962 Human pol |
| 16 | 34.2 | 29.5 | 1777 | 6 ADD18987 | ADD18987 Human d1s |
| 17 | 34.2 | 29.5 | 201143 | 6 ABK83568 | ABK83568 Human DNA |
| 18 | 30.4 | 26.2 | 823 | 6 ADE25658 | ADE25658 Human CDN |
| 19 | 30.4 | 26.2 | 1055 | 6 AAD24392 | AAD24392 Human RNA |
| 20 | 30.4 | 26.2 | 1427 | 4 AAK73067 | AAK73067 Human imm |
| 21 | 30.4 | 26.2 | 1526 | 9 ADD18895 | ADD18895 Human d1s |
| 22 | 30.4 | 26.2 | 1567 | 4 AAI93080 | AAI93080 Human pol |
| 23 | 30.4 | 26.2 | 1638 | 6 ABQ61187 | ABQ61187 Unidentif |

| | | | | | |
|----|------|------|--------|---------------|--------------------|
| 24 | 30.2 | 26.0 | 945 | 6 ABQ69695 | ABQ69695 Listeria |
| 25 | 30.2 | 26.0 | 945 | 6 ABQ69758 | ABQ69758 Listeria |
| 26 | 30.2 | 26.0 | 966 | 6 ABQ67740 | ABQ67740 Listeria |
| 27 | 30.2 | 26.0 | 966 | 6 ABQ67741 | ABQ67741 Listeria |
| 28 | 30.2 | 26.0 | 84707 | 6 ABQ67196_5 | Continuation (7 of |
| 29 | 30.2 | 26.0 | 110000 | 6 ABQ67196_1 | Continuation (2 of |
| 30 | 30.2 | 26.0 | 110000 | 6 ABQ69245_24 | Continuation (25 o |
| 31 | 30.2 | 26.0 | 110000 | 6 ABQ69245_26 | Continuation (27 o |
| 32 | 29 | 25.0 | 7113 | 4 ABA09050 | ABA09050 Human sec |
| 33 | 29 | 25.0 | 7113 | 4 ABA09050 | ABA09050 Human sec |
| 34 | 29 | 25.0 | 7287 | 4 AAK51481 | AAK51481 Human pol |
| 35 | 28.8 | 24.8 | 274 | 5 ABV89929 | ABV89929 Human pro |
| 36 | 28.6 | 24.7 | 15181 | 4 AAK65588 | AAK65588 Human imm |
| 37 | 28.6 | 24.7 | 15185 | 4 AAK65589 | AAK65589 Human imm |
| 38 | 28.6 | 24.7 | 68230 | 8 ADA66349 | ADA66349 Mouse Ppp |
| 39 | 28.6 | 24.7 | 68233 | 8 ADA03065 | ADA03065 Mouse Ppp |
| 40 | 28.6 | 24.7 | 68233 | 9 ADB72803 | ADB72803 Mouse Ppp |
| 41 | 28.2 | 24.3 | 743 | 4 AAI94932 | AAI94932 Mouse neu |
| 42 | 28 | 24.1 | 28 | 3 AAC5332 | AAC5332 Human act |
| 43 | 28 | 24.1 | 2000 | 7 ADA71431 | ADA71431 Rice gene |
| 44 | 27.8 | 24.0 | 261 | 4 AAS24582 | AAS24582 Human ova |
| 45 | 27.8 | 24.0 | 390 | 5 AAH83200 | AAH83200 Human ova |

ALIGNMENTS

| | |
|----------|--|
| RESULT 1 | |
| AAC5318 | AAC5318 standard; DNA; 116 BP. |
| AC | AAC5318; |
| DT | 05-FEB-2001 (first entry) |
| XX | Human activation-induced cytidine deaminase exon 4 SEQ ID NO:14. |
| DE | |
| XX | |
| KW | Activation-induced cytidine deaminase; A1D; cytidine deaminase; |
| KW | immune related disease; allergy; allergic disease; antiallergic; |
| KW | antianemic; antiaesthetic; ophthalmological; anti-HIV; dermatological; |
| KW | gene therapy; B cell associated immune system disorder; food allergy; |
| KW | immunodeficiency disease; immunoglobulin A deficiency disease; asthma; |
| KW | IgA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis; |
| KW | drug allergy; allergic rhinitis; Rosen disease; Pigeon disease; AIDS; |
| KW | ataxia telangiectasia; common variable immunodeficiency disorder; |
| KW | major histocompatibility class II deficiency disease; |
| KW | auto immunodeficiency syndrome; IgG subclass selection disorder; ds. |
| OS | Homo sapiens. |
| XX | |
| PN | WO200058480-A1. |
| XX | |
| PD | 05-OCT-2000. |
| XX | |
| PF | 28-MAR-2000; 2000WO-JP001918. |
| XX | |
| PR | 29-MAR-1999; 99JP-00087192. |
| PR | 24-JUN-1999; 99JP-00178999. |
| PR | 27-DEC-1999; 99JP-00371382. |
| XX | |
| PA | (NTSB) JAPAN TOBACCO INC. |
| PA | (HONJ) HONJO T. |
| PI | Honjo T, Muramatsu M; |
| DR | WPI, 2000-611715/58. |
| XX | |
| PT | Nucleic acid encoding activation induced cytidine deaminase, useful as a |
| PT | target for drug development for immune-related diseases including |
| PT | allergies. |
| XX | |
| FS | Claim 18; Page 151; 174pp; Japanese. |
| XX | |

CC The present invention describes an activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has anti-allergic, anti-nausea,
 CC antidiabetic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
 CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, Digeorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, MHC (major
 CC histocompatibility class) class II deficiency disease, AIDS (auto
 CC immunodeficiency syndrome), elevated Igg disorder, and Igg subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders. The present sequence represents the
 CC exon 4 genomic DNA sequence of human AID

SQ Sequence 116 BP; 32 A; 25 C; 24 G; 35 T; 0 U; 0 Other;

Query Match 100.0%; Score 116; DB 3; Length 116;

Best Local Similarity 100.0%; Pred. No. 3.6e-29; Mismatches 0; Gaps 0;

Matches 116; Conservative 0; Indels 0; Gaps 0;

1 ATTATTTTACTGCTGGAATCTTTGTAGAAAACCAAGAACTTCAAGCTGGG 60

1 ATTATTTTACTGCTGGAATCTTTGTAGAAAACCAAGAACTTCAAGCTGGG 60

61 AAGGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTGGGCGATCCTTTG 116

61 AAGGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTGGGCGATCCTTTG 116

RESULT 2

AAC55314

XX AAC55314; DNA; 6564 BP.

AC AAC55314;

DE 05-FEB-2001 (first entry)

XX Human activation-induced cytidine deaminase genomic DNA SEQ ID NO:10.

XX Activation-induced cytidine deaminase; AID; cytidine deaminase;

KM immune related disease; allergy; allergic disease; anti-allergic;

KM antinausea; antidiabetic; ophthalmological; anti-HIV; dermatological;

KM gene therapy; B cell associated immune system disorder; food allergy;

KM immunodeficiency disease; immunoglobulin A deficiency disease; asthma;

KM IGA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;

KM drug allergy; allergic rhinitis; Rosen disease; Digeorge disease; AIDS;

KM ataxia telangiectasia; common variable immunodeficiency disorder;

KM major histocompatibility class II deficiency disease;

KM auto immunodeficiency syndrome; Igg subclass selection disorder; ds.

XX Homo sapiens.

OS Homo sapiens.

PN WO200058480-A1.

PD 05-OCT-2000.

PF 26-MAR-2000; 2000WO-JP001918.

PR 29-MAR-1999; 99JP-00087192.

PR 24-JUN-1999; 99JP-00178999.

PR 27-DEC-1999; 99JP-00371382.

XX (NISR) JAPAN TOBACCO INC.

PA (HONJ/) HONJO T.

XX Honjo T, Muramatsu M;

PI WPI, 2000-611715/58.

XX

FT Nucleic acid encoding activation induced cytidine deaminase, useful as a
 PT target for drug development for immune-related diseases including
 PT allergies.

XX Claim 17; Page 145-150; 174pp; Japanese.

CC The present invention describes an activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has anti-allergic, anti-nausea,
 CC antidiabetic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
 CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, Digeorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, MHC (major
 CC histocompatibility class) class II deficiency disease, AIDS (auto
 CC immunodeficiency syndrome), elevated Igg disorder, and Igg subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders. The present sequence represents a
 CC genomic DNA sequence of human AID

SQ Sequence 6564 BP; 1909 A; 1358 C; 1383 G; 1914 T; 0 U; 0 Other;

Query Match 100.0%; Score 116; DB 3; Length 6564;

Best Local Similarity 100.0%; Pred. No. 1.3e-28; Mismatches 0; Gaps 0;

Matches 116; Conservative 0; Indels 0; Gaps 0;

1 ATTATTTTACTGCTGGAATCTTTGTAGAAAACCAAGAACTTCAAGCTGGG 60

1 ATTATTTTACTGCTGGAATCTTTGTAGAAAACCAAGAACTTCAAGCTGGG 60

61 AAGGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTGGGCGATCCTTTG 116

61 AAGGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTGGGCGATCCTTTG 116

RESULT 3

AAC55339

XX AAC55339; DNA; 11204 BP.

AC AAC55339;

DE 05-FEB-2001 (first entry)

XX Human activation-induced cytidine deaminase genomic DNA SEQ ID NO:35.

XX Activation-induced cytidine deaminase; AID; cytidine deaminase;

KM immune related disease; allergy; allergic disease; anti-allergic;

KM antinausea; antidiabetic; ophthalmological; anti-HIV; dermatological;

KM gene therapy; B cell associated immune system disorder; food allergy;

KM immunodeficiency disease; immunoglobulin A deficiency disease; asthma;

KM IGA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;

KM drug allergy; allergic rhinitis; Rosen disease; Digeorge disease; AIDS;

KM ataxia telangiectasia; common variable immunodeficiency disorder;

KM major histocompatibility class II deficiency disease;

KM auto immunodeficiency syndrome; Igg subclass selection disorder; ds.

XX Homo sapiens.

OS Homo sapiens.

PN WO200058480-A1.

PD 05-OCT-2000.

PF 26-MAR-2000; 2000WO-JP001918.

PR 29-MAR-1999; 99JP-00087192.

PR 24-JUN-1999; 99JP-00178999.

PR 27-DEC-1999; 99JP-00371382.

XX (NISR) JAPAN TOBACCO INC.

| | | |
|----|---|--|
| XX | (HONU//) HONJO T. | |
| PI | Honjo T, Muramatsu M; | |
| XX | WPI; 2000-611715/58. | |
| XX | | |
| PT | Nucleic acid encoding activation induced cytidine deaminase, useful as a | |
| PT | target for drug development for immune-related diseases including | |
| PT | allergies. | |
| PS | Claim 17; Page 163-170; 174pp; Japanese. | |
| XX | | |
| CC | The present invention describes an activation-induced cytidine deaminase | |
| CC | (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has | |
| CC | cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic, | |
| CC | antiaesthetic, ophthalmological, anti-HIV and dermatological activities, | |
| CC | and can be used in gene therapy. AID polynucleotides are useful in | |
| CC | methods for identifying drugs for the treatment of B cell associated | |
| CC | immune system disorders, immunodeficiency diseases and allergies, such as | |
| CC | immunoglobulin A (IGA) deficiency disease, Iga nephritis, gamma- | |
| CC | globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy, | |
| CC | drug allergy, allergic rhinitis, Rosen disease, Digorge disease, ataxia | |
| CC | telangiectasia, common variable immunodeficiency disorder, MHC (major | |
| CC | histocompatibility class II deficiency disease, AIDS (auto | |
| CC | immunodeficiency syndrome), elevated Ige disorder, and Igg subclass | |
| CC | selection disorder. The DNA sequences encoding AID may be used for gene | |
| CC | therapy and the antibodies to the AID protein may be used for diagnosis | |
| CC | and treatment of these disorders. The present sequence represents a | |
| XX | genomic DNA sequence of human AID | |
| XX | | |
| SO | Sequence 11204 BP; 3305 A; 2273 G; 3253 T; 0 U; 0 Other; | |
| | | |
| | Query Match | 100.0%; Score 116; DB 3; Length 11204; |
| | Best Local Similarity | 100.0%; Pred. No. 1.5e-26; |
| | Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| OY | 1 ATATATTTTACTGCTCGAATACTTTTGTAGAAAACCAAGAAAGAACTTCAAAGCTGGG 60 | |
| Db | 8371 ATATATTTTACTGCTCGAATACTTTTGTAGAAAACCAAGAAAGAACTTCAAAGCTGGG 8430 | |
| OY | 61 AAGGCTCATGAATAATTCAGTTCGCTCTCCAGACAGCTTCGGGCGATCCTTTTG 116 | |
| Db | 8431 AAGGCTCATGAATAATTCAGTTCGCTCTCTCCAGACAGCTTCGGGCGATCCTTTTG 8486 | |
| | | |
| | RESULT 4 | |
| AC | ABS73286 standard; DNA; 11204 BP. | |
| XX | | |
| XX | ABS73286; | |
| XX | | |
| XX | 04-DEC-2002 (first entry) | |
| XX | | |
| DE | DNA encoding human translocation del(12p) protein #1. | |
| XX | | |
| KW | Chromosome aberration; oncogenic fusion protein; cancer; | |
| KW | proliferative disease; cellular protein isoform; heat shock protein 90; | |
| KW | HSP-90; rheumatoid arthritis; cancer; hematopoietic disorder; | |
| KW | T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML; | |
| KW | acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML; | |
| KW | acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour; | |
| KW | papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma; | |
| KW | rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds. | |
| XX | | |
| XX | Homo sapiens. | |
| XX | | |
| XX | WO200269900-A2. | |
| XX | | |
| XX | 12-SEP-2002. | |
| XX | | |
| XX | 01-MAR-2002; 2002MO-US006518. | |
| XX | | |
| XX | 01-MAR-2001; 2001US-0272751P. | |

| | | | |
|----------|----------|---|--|
| XX | PA | (CONF-) | CONFORMA THERAPEUTICS CORP. |
| XX | P1 | Fritz LC, Burrows EJ; | |
| XX | XX | WPI; 2002-698710/75. | |
| DR | P-PSDB | ABG95082. | |
| XX | PT | Treating genetically-defined disease associated with chromosomal | |
| XX | PT | aberrations yielding oncogenic fusion proteins, e.g. cell proliferative | |
| XX | PT | diseases, involves administering an inhibitor of heat shock protein 90. | |
| PS | PS | Disclosure; Page 242-245; 389pp; English. | |
| XX | CC | The invention describes a method of treating genetically-defined disease | |
| XX | CC | associated with chromosomal aberrations yielding oncogenic fusion | |
| XX | CC | proteins (I), treating cancerous cells containing (I) in a heterogeneous | |
| XX | CC | cell population, treating proliferative diseases associated with mutant | |
| XX | CC | protein or cellular protein isoforms (II) dependent on heat shock protein | |
| XX | CC | (HSP)-90, or selectively treating cells expressing (II) involving | |
| XX | CC | administering HSP90-inhibitor. The method is useful for treating | |
| XX | CC | genetically-defined disease with chromosomal aberration yielding | |
| XX | CC | oncogenic fusion protein, treating cancerous cells containing fusion | |
| XX | CC | protein in heterogeneous cell population, treating proliferative disease | |
| XX | CC | (e.g. rheumatoid arthritis or cancer) associated with mutant protein or | |
| XX | CC | cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g. | |
| XX | CC | p53), or selectively treating cells expressing mutant protein or cellular | |
| XX | CC | protein isoform in a patient heterozygous for (II). The method is useful | |
| XX | CC | for treating a disease e.g. hematopoietic disorder such as T or B cell | |
| XX | CC | lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CML, | |
| XX | CC | or a disease characterised by a solid tumour such as papillary thyroid | |
| XX | CC | carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and | |
| XX | CC | synovial sarcoma. The method is also useful for treating viral | |
| XX | CC | infections. This represents the DNA sequence of a chromosome aberration | |
| XX | XX | Sequence 11204 BP; 3305 A; 2273 C; 2373 G; 3253 T; 0 U; 0 Other; | |
| XX | XX | Query Match 100.0%; Score 116; DB 6; Length 11204; | |
| XX | XX | Best Local Similarity 100.0%; Pred. No. 1.5e-28; | |
| XX | XX | Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| QY | 1 | ATTATTTTCTGCTGGAAATCTTTTGAGAAACACGAAAGACTTCAAGGCTGG 60 | |
| DB | 8371 | ATTATTTTCTGCTGGAAATCTTTTGAGAAACACGAAAGACTTCAAGGCTGG 8430 | |
| QY | 61 | AAGGCTGCATGAAATTCAGTTCGCTCTCCAGACAGCTTCGGCGATCCTTTG 116 | |
| DB | 8431 | AAGGCTGCATGAAATTCAGTTCGCTCTCCAGACAGCTTCGGCGATCCTTTG 8486 | |
| RESULT 5 | | | |
| AC | ADB88952 | ADB88952 standard; DNA; 597 BP. | |
| XX | XX | ADB88952; | |
| XX | DT | 04-DEC-2003 (first entry) | |
| XX | DE | AID gene as substrate used in AID mutagenic method. | |
| XX | XX | des. cytostatic; virucide; antidote; mutation; | |
| XX | KW | activation-induced cytidine deaminase; cell phenotype; | |
| XX | KW | monoclonal antibody; antigen; cross-reactivity; tumour; | |
| XX | KW | gem warfare agent; toxin; V region. | |
| OS | XX | Homo sapiens. | |
| XX | XX | Key | Location/Qualifiers |
| FT | FT | mutation | replace(9,A) |
| FT | FT | mutation | /*tag= a |
| FT | FT | mutation | /note= "mutation in P1-5 hybridoma cell" |
| FT | FT | mutation | replace(12,A) |
| FT | FT | mutation | /*tag= b |

```

FT mutation /note= "mutation in Ramos cell"
FT replace(14,C)
FT /*tag= c
FT /note= "mutation in CHO cell"
FT mutation replace(61,T)
FT /*tag= d
FT /note= "mutation in P1-5 hybridoma cell"
FT mutation replace(66,A)
FT /*tag= e
FT /note= "mutation in P1-5 hybridoma cell"
FT mutation replace(73,T)
FT /*tag= f
FT /note= "mutation in CHO cell"
FT mutation replace(79,G)
FT /*tag= g
FT /note= "mutation in Ramos cell"
FT mutation replace(85,G)
FT /*tag= h
FT /note= "mutation in P1-5 hybridoma cell"
FT mutation replace(112,C)
FT /*tag= i
FT /note= "mutation in Ramos cell"
FT mutation replace(118,G)
FT /*tag= j
FT /note= "mutation in P1-5 hybridoma cell"
FT mutation replace(156,T)
FT /*tag= k
FT /note= "mutation in P1-5 hybridoma cell"
FT mutation replace(161,A)
FT /*tag= l
FT /note= "mutation in P1-5 hybridoma cell"
FT mutation replace(165,T,G)
FT /*tag= m
FT /note= "G mutation in Ramos cell, T mutation in CHO cell"
FT mutation replace(178,T)
FT /*tag= n
FT /note= "mutation in Ramos cell"
FT mutation replace(206,G)
FT /*tag= o
FT /note= "mutation in CHO cell"
FT mutation replace(209,C)
FT /*tag= p
FT /note= "mutation in Ramos cell"
FT mutation replace(218,T)
FT /*tag= q
FT /note= "mutation in Ramos cell"
FT mutation replace(250,C)
FT /*tag= r
FT /note= "mutation in Ramos cell"
FT mutation replace(261,G)
FT /*tag= s
FT /note= "mutation in Ramos cell"
FT mutation replace(289,A)
FT /*tag= t
FT /note= "mutation in Ramos cell"
FT mutation replace(303,T)
FT /*tag= u
FT /note= "mutation in Ramos cell"
FT mutation replace(329,T)
FT /*tag= v
FT /note= "mutation in Ramos cell"
FT mutation replace(350,G,T)
FT /*tag= w
FT /note= "T mutation in Ramos cell, G mutation in CHO cell"
FT mutation replace(356,A)
FT /*tag= x
FT /note= "A mutation in Ramos cell, A mutation in CHO cell"
FT mutation replace(357,G)
FT /*tag= y
FT /note= "mutation in Ramos cell"
FT mutation replace(367,T, A)
FT /*tag= z
FT /note= "mutation in P1-5 hybridoma cell"

```

```

FT mutation replace(375,A)
FT /*tag= aa
FT /note= "mutation in CHO cell"
FT mutation replace(402,A)
FT /*tag= ab
FT /note= "mutation in P1-5 hybridoma cell"
FT mutation replace(410,T)
FT /*tag= ac
FT /note= "mutation in P1-5 hybridoma cell"
FT mutation replace(418,T)
FT /*tag= ad
FT /note= "mutation in P1-5 hybridoma cell"
FT mutation replace(420,T)
FT /*tag= ae
FT /note= "mutation in P1-5 hybridoma cell"
FT mutation replace(421,A)
FT /*tag= af
FT /note= "mutation in CHO cell"
FT mutation replace(430,G)
FT /*tag= ag
FT /note= "mutation in Ramos cell"
FT mutation replace(440,A)
FT /*tag= ah
FT /note= "mutation in CHO cell"
FT mutation replace(449,T)
FT /*tag= ai
FT /note= "mutation in P1-5 hybridoma cell"
FT mutation replace(495,A)
FT /*tag= aj
FT /note= "mutation in CHO cell"
FT mutation replace(517,C)
FT /*tag= ak
FT /note= "mutation in Ramos cell"
FT mutation replace(526,T, A)
FT /*tag= al
FT /note= "T mutation in P1-5 hybridoma cell, A mutation in Ramos cell"
FT mutation replace(533,A)
FT /*tag= am
FT /note= "mutation in P1-5 hybridoma cell"
FT mutation replace(538,A)
FT /*tag= an
FT /note= "mutation in CHO cell"
FT mutation replace(549,A)
FT /*tag= ao
FT /note= "mutation in CHO cell"
FT mutation replace(568,T)
FT /*tag= ap
FT /note= "mutation in Ramos cell"
FT mutation replace(570,T)
FT /*tag= aq
FT /note= "mutation in CHO cell"

WO2003061363-A2.
31-JUL-2003.
15-JAN-2003; 2003WO-US001149.
17-JAN-2002; 2002US-0350269P.
(YESH ) UNIV YESHIVA EITNSTEIN COLLEGE.
Martin A, Scharff MD;
WPI, 2003-747983/70.

Induction of mutations in gene expressed in eukaryotic cell useful e.g.
to produce antibodies with greater affinity or specificity for antigens
e.g. human pathogens by expressing an activation-induced cytidine
deaminase gene in cell.

Example 4, Fig 6; 73pp; English.

```

XX The invention relates to the induction of mutations in a gene expressed
 CC in a eukaryotic cell, where the gene is operably linked to a promoter and
 CC within 2 kb of promoter, by expressing a transgenic activation-induced
 CC cytidine deaminase (AID) gene in the cell. The method is useful for
 CC producing mutated genes and proteins, determining the effect of a
 CC mutation on a protein or cell phenotype, and producing genes and proteins
 CC with altered properties. It is especially useful for producing antibodies
 CC with altered affinities or specificities for an antigen, or
 CC increased/decreased cross-reactivity for a second antigen, e.g. to
 CC produce antibodies with greater affinity/specificity useful
 CC therapeutically to treat tumours, viruses such as Ebola and Lassa Fever
 CC or against germ warfare agents, toxins (e.g. ricin). The method may
 CC similarly be used to alter affinity, specificity or cross-reactivity of a
 CC monoclonal antibody. This sequence represents the AID gene used in the
 CC method of the invention to induce mutations in the sequence.

XX
 SQ Sequence 597 BP, 128 A; 164 C; 155 G; 150 T; 0 U; 0 Other;

Query Match 98.6%; Score 114.4; DB 8; Length 597;
 Best Local Similarity 99.1%; Pred. No. 2.1e-28;
 Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATATATTTTACTGCTGGAATACCTTTGTAGAAAACGAAAGAACTTCAAGCCTGGG 60
 DB 428 ATTATTTTACTGCTGGAATACCTTTGTAGAAAACGAAAGAACTTCAAGCCTGGG 487
 QY 61 AAGGCGTCATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 116
 DB 488 AAGGCGTCATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 543

RESULT 6
 ABX05468
 ID ABX05468 standard; cDNA; 1543 BP.

XX
 AC ABX05468;
 XX
 DT 17-JAN-2003 (first entry)
 XX
 DE Human novel polynucleotide #483.

XX Human; gene; ss; genetic disorder; gene mapping; medical imaging; cancer;
 KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
 KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
 KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;
 KW fungal infection; bacterial infection; autoimmune disease; diabetes;
 KW atopic dermatitis.

XX Homo sapiens.

XX WO200274961-A1.

XX 26-SEP-2002.

XX 14-MAR-2002; 2002MO-US005109.

XX 15-MAR-2001; 2001US-00810173.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Drmanac RT;

XX WPI; 2003-040556/03.
 DR P-PSDB; ABU00390.

XX New isolated polypeptides and polynucleotides, useful for preventing,
 PT treating or ameliorating medical conditions, such as cancer,
 PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
 PT disorders, and infections.

PS Claim 1; SEQ ID NO 483; 235pp; English.

XX The invention relates to human polynucleotides and the polypeptides they
 CC encode. The polynucleotides and polypeptides are useful in diagnostics,
 CC forensics, gene mapping, medical imaging, identification of mutations,
 CC responsible for genetic disorders or other traits, assessing biodiversity
 CC and producing many other types of data and products dependent on DNA and
 CC amino acid sequences. They are also useful for preventing, treating or
 CC ameliorating medical conditions, such as cancer, neurodegenerative
 CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
 CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
 CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
 CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
 CC Sequences ABX04986-ABX05511 represent human polynucleotides of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification but is based on sequence information supplied
 CC by the European Patent Office

XX
 SQ Sequence 1543 BP, 428 A; 334 C; 352 G; 429 T; 0 U; 0 Other;

Query Match 98.6%; Score 114.4; DB 7; Length 1543;
 Best Local Similarity 99.1%; Pred. No. 2.8e-28;
 Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATATATTTTACTGCTGGAATACCTTTGTAGAAAACGAAAGAACTTCAAGCCTGGG 60
 DB 505 ATTATTTTACTGCTGGAATACCTTTGTAGAAAACGAAAGAACTTCAAGCCTGGG 564
 QY 61 AAGGCGTCATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 116
 DB 565 AAGGCGTCATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 620

RESULT 7
 ABS73287
 ID ABS73287 standard; DNA; 2791 BP.

XX
 AC ABS73287;
 XX
 DT 04-DEC-2002 (first entry)
 XX
 DE DNA encoding human translocation del(12p) protein #2.

XX Chromosome aberration; oncogenic fusion protein; cancer;
 KW proliferative disease; cellular protein isoform; heat shock protein 90;
 KW HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
 KW T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
 KW acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
 KW acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
 KW papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
 KW rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.

XX Homo sapiens.

XX WO200269900-A2.

XX 12-SEP-2002.

XX 01-MAR-2002; 2002MO-US006518.

XX 01-MAR-2001; 2001US-0272751P.

XX (CONF-) CONFORMA THERAPEUTICS CORP.

XX Fritz LC, Burrows FU;

XX WPI; 2002-698710/75.
 DR P-PSDB; ABG95083.

XX Treating genetically-defined disease associated with chromosomal
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
 PT diseases, involves administering an inhibitor of heat shock protein 90.

PS Disclosure; Page 246-247; 389pp; English.

XX The invention describes a method of treating genetically-defined disease
CC associated with chromosomal aberrations yielding oncogenic fusion
CC proteins (I), treating cancerous cells containing (I) in a heterogeneous
CC cell population, treating proliferative diseases associated with mutant
CC protein or cellular protein isoforms (II) dependent on heat shock protein
CC (HSP)-90, or selectively treating cells expressing (II) involving
CC administering HSP90-inhibitor. The method is useful for treating
CC genetically-defined disease with chromosomal aberration yielding
CC oncogenic fusion protein, treating cancerous cells containing fusion
CC protein in heterogeneous cell population, treating proliferative disease
CC (e.g., rheumatoid arthritis or cancer) associated with mutant protein or
CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.,
CC p53), or selectively treating cells expressing mutant protein or cellular
CC protein isoform in a patient heterozygous for (II). The method is useful
CC for treating a disease e.g., hematopoietic disorder such as T or B cell
CC lymphoma, chronic myeloid leukemia (CML), APL, ALL, AML, NHL and CML,
CC or a disease characterized by a solid tumour such as papillary thyroid
CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
CC synovial sarcoma. The method is also useful for treating viral
CC infections. This represents the DNA sequence of a chromosome aberration
XX

SO Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;

Query Match 98.6%; Score 114.4; DB 6; Length 2791;
Best Local Similarity 99.1%; Pred. No. 3,4e-28;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAAATCTTTGTAGAAAACCGAAGAACTTTCAAGCCTGG 60
DB 504 ATTATTTTACTGCTGGAAATCTTTGTAGAAAACCGAAGAACTTTCAAGCCTGG 563
61 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCCATCCTTTG 116
DB 564 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCCATCCTTTG 619

RESULT 8
ABS73288
ID ABS73288 standard; DNA; 2791 BP.

XX ABS73288;
XX
XX 04-DEC-2002 (first entry)
XX
XX
DE DNA encoding human translocation del(12p) protein #3.
XX
XX Chromosome aberration; oncogenic fusion protein; cancer;
XX proliferative diseases; cellular protein isoform; heat shock protein 90;
XX HSP-90; rheumatoid arthritis; cancer; hematopoietic disorder;
XX T cell lymphoma; B cell lymphoma; chronic myeloid leukemia; CML;
XX acute myeloid leukemia; AML; chronic myelomonocytic leukemia; CMML;
XX acute lymphoblastic leukemia; ALL; APL; NHL; solid tumour;
XX papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
XX rhabdomyosarcoma; synovial sarcoma; viral infection; gene; de.
XX
XX Homo sapiens.
XX
XX WO200269900-A2.
XX
XX 12-SEP-2002.
XX
XX 01-MAR-2002; 2002WO-US006518.
XX
XX 01-MAR-2001; 2001US-0272751P.
XX
XX (CONF-) CONFORMA THERAPEUTICS CORP.
XX
XX Fritz LC, Burrows FU;
XX
XX WPI; 2002-698710/75.
XX
XX P-PDB; ABG95084.
XX

XX Treating genetically-defined disease associated with chromosomal
PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
PT diseases, involves administering an inhibitor of heat shock protein 90.
XX
XX Disclosure; Page 246-249; 389pp; English.

XX The invention describes a method of treating genetically-defined disease
CC associated with chromosomal aberrations yielding oncogenic fusion
CC proteins (I), treating cancerous cells containing (I) in a heterogeneous
CC cell population, treating proliferative diseases associated with mutant
CC protein or cellular protein isoforms (II) dependent on heat shock protein
CC (HSP)-90, or selectively treating cells expressing (II) involving
CC administering HSP90-inhibitor. The method is useful for treating
CC genetically-defined disease with chromosomal aberration yielding
CC oncogenic fusion protein, treating cancerous cells containing fusion
CC protein in heterogeneous cell population, treating proliferative disease
CC (e.g., rheumatoid arthritis or cancer) associated with mutant protein or
CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.,
CC p53), or selectively treating cells expressing mutant protein or cellular
CC protein isoform in a patient heterozygous for (II). The method is useful
CC for treating a disease e.g., hematopoietic disorder such as T or B cell
CC lymphoma, chronic myeloid leukemia (CML), APL, ALL, AML, NHL and CMML,
CC or a disease characterized by a solid tumour such as papillary thyroid
CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
CC synovial sarcoma. The method is also useful for treating viral
CC infections. This represents the DNA sequence of a chromosome aberration
XX

SO Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;

Query Match 98.6%; Score 114.4; DB 6; Length 2791;
Best Local Similarity 99.1%; Pred. No. 3,4e-28;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAAATCTTTGTAGAAAACCGAAGAACTTTCAAGCCTGG 60
DB 504 ATTATTTTACTGCTGGAAATCTTTGTAGAAAACCGAAGAACTTTCAAGCCTGG 563
61 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCCATCCTTTG 116
DB 564 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCCATCCTTTG 619

RESULT 9
AAC55312
ID AAC55312 standard; cDNA; 2818 BP.

XX AAC55312;
XX
XX 05-FEB-2001 (first entry)
XX
XX
DE Human activation-induced cytidine deaminase encoding cDNA SEQ ID NO:7.
XX
XX Activation-induced cytidine deaminase; AID; cytidine deaminase;
XX immune related disease; allergy; allergic disease; anti-allergic;
XX antianemic; antisthmatic; ophthalmological; anti-HIV; dermatological;
XX gene therapy; B cell associated immune system disorder; food allergy;
XX immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
XX IgA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;
XX drug allergy; allergic rhinitis; Rosen disease; Digorge disease; AIDS;
XX ataxia telangiectasia; common variable immunodeficiency disorder;
XX major histocompatibility class II deficiency disease;
XX auto immunodeficiency syndrome; IgG subclass selection disorder; ss.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 80..676
XX /*tag= a
XX /product= "activation-induced cytidine deaminase"
XX
XX WO200058480-A1.
XX

PD 05-OCT-2000.
 XX
 PF 28-MAR-2000; 2000WO-JP001918.
 XX
 PR 29-MAR-1999; 99JP-00087192.
 PR 24-JUN-1999; 99JP-00178999.
 PR 27-DEC-1999; 99JP-00371382.
 XX
 PA (NIBS) JAPAN TOBACCO INC.
 PA (HONJ/) HONJO T.
 XX
 PI Horjo T, Muramatsu M;
 XX
 PF WPI; 2000-611715/58.
 DR P-PSDB; AAB24198.
 XX
 PT Nucleic acid encoding activation induced cytidine deaminase, useful as a
 PT target for drug development for immune-related diseases including
 PT allergies.
 XX
 PS Claim 3; Page 135-139; 174pp; Japanese.
 XX
 CC The present sequence encodes human activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has anti-allergic, anti-inflammatory,
 CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
 CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, DiGeorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, MHC (major
 CC histocompatibility class II deficiency disease, AIDS (auto
 CC immunodeficiency syndrome), elevated Igb disorder, and Igg subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders
 XX
 SQ Sequence 2818 BP; 868 A; 548 C; 626 G; 776 T; 0 U; 0 Other;
 Query Match 98.6%; Score 114.4; DB 3; Length 2818;
 Best Local Similarity 99.1%; Pred. No. 3.4e-28;
 Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATTAATTTTACTGCTGGAATACCTTTGTAGAAAACCAAGAAAGCTTTCAAGCCTGGG 60
 DB 507 ATTAATTTTACTGCTGGAATACCTTTGTAGAAAACCAAGAAAGCTTTCAAGCCTGGG 566
 QY 61 AAGGCTGATGAATAATTCAGTTGCTCTCTCCAGACAGCTTCGGCGATCCTTTTG 116
 DB 567 AAGGCTGATGAATAATTCAGTTGCTCTCTCCAGACAGCTTCGGCGATCCTTTTG 622
 XX
 RESULT 10
 ID AACS5307 standard; cDNA; 2440 BP.
 XX
 AC AACS5307;
 XX
 DT 05-FEB-2001 (first entry)
 XX
 DE Mouse activation-induced cytidine deaminase encoding cDNA SEQ ID NO:1.
 XX
 KW Activation-induced cytidine deaminase; AID; cytidine deaminase;
 KW immune related disease; allergy; allergic disease; anti-allergic;
 KW anti-inflammatory; antiasthmatic; ophthalmological; anti-HIV; dermatological;
 KW gene therapy; B cell associated immune system disorder; food allergy;
 KW immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
 KW IGA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;
 KW drug allergy; allergic rhinitis; Rosen disease; DiGeorge disease; AIDS;
 KW ataxia telangiectasia; common variable immunodeficiency disorder;
 KW major histocompatibility class II deficiency disease;

KW auto immunodeficiency syndrome; Igg subclass selection disorder; ss.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT CDS 93..689
 FT /*tag= a
 FT /product= "activation-induced cytidine deaminase"
 XX
 PN WO200058480-A1.
 XX
 PD 05-OCT-2000.
 XX
 PF 28-MAR-2000; 2000WO-JP001918.
 XX
 PR 29-MAR-1999; 99JP-00087192.
 PR 24-JUN-1999; 99JP-00178999.
 PR 27-DEC-1999; 99JP-00371382.
 XX
 PA (NIBS) JAPAN TOBACCO INC.
 PA (HONJ/) HONJO T.
 XX
 PI Horjo T, Muramatsu M;
 XX
 DR WPI; 2000-611715/58.
 DR P-PSDB; AAB24197.
 XX
 PT Nucleic acid encoding activation induced cytidine deaminase, useful as a
 PT target for drug development for immune-related diseases including
 PT allergies.
 XX
 PS Claim 3; Page 126-130; 174pp; Japanese.
 XX
 CC The present sequence encodes mouse activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has anti-allergic, anti-inflammatory,
 CC antiasthmatic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
 CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, DiGeorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, MHC (major
 CC histocompatibility class II deficiency disease, AIDS (auto
 CC immunodeficiency syndrome), elevated Igb disorder, and Igg subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders
 XX
 SQ Sequence 2440 BP; 706 A; 546 C; 551 G; 636 T; 0 U; 1 Other;
 Query Match 83.4%; Score 96.8; DB 3; Length 2440;
 Best Local Similarity 89.7%; Pred. No. 2.7e-22;
 Matches 104; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 1 ATTAATTTTACTGCTGGAATACCTTTGTAGAAAACCAAGAAAGCTTTCAAGCCTGGG 60
 DB 520 ACTAATTTTACTGCTGGAATACCTTTGTAGAAAACCAAGAAAGCTTTCAAGCCTGGG 579
 QY 61 AAGGCTGATGAATAATTCAGTTGCTCTCTCCAGACAGCTTCGGCGATCCTTTTG 116
 DB 580 AAGGCTGATGAATAATTCAGTTGCTCTCTCCAGACAGCTTCGGCGATCCTTTTG 635
 XX
 RESULT 11
 ABN96785
 ID ABN96785 standard; DNA; 1348 BP.
 XX
 AC ABN96785;
 XX
 DT 13-AUG-2002 (first entry)
 XX

DE Gene #3283 used to diagnose liver cancer.
XX
XX
KM Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KM metastatic liver tumor; cytostatic; expression profile; disease state;
KM disease progression; drug toxicity; drug efficacy; drug metabolism.
OS
XX Homo sapiens.
XX
XX MO200229103-A2.
XX
XX 11-APR-2002.
XX
XX PD 02-OCT-2001; 2001MO-US030589.
XX
XX PF 02-OCT-2000; 2000US-0237054P.
XX
XX FR 02-OCT-2000; 2000US-0237054P.
XX
XX PA (GENE-) GENE LOGIC INC.
XX
XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.
XX
XX DR
XX PT Diagnosing and detecting the progression of liver cancer, hepatocellular
XX PT carcinoma or metastatic liver tumor in a patient, involves detecting the
XX PT level of expression of two or more genes in a liver tissue sample.
XX
XX PS Claim 1; SEQ ID NO 3283; 298bp; English.
XX
XX CC The invention relates to a novel method for diagnosing and detecting the
XX CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX CC tumor in a patient, and differentiating metastatic liver cancer from
XX CC hepatocellular carcinoma in a patient, involving detecting the level of
XX CC expression of two or more genes represented in ABN93503-ABN97455 in a
XX CC tissue sample. The method of the invention has hepatotropic, and
XX CC cyrostatic activity. The method is useful for diagnosing and detecting
XX CC the progression of liver cancer, hepatocellular carcinoma and metastatic
XX CC liver carcinoma in a patient. The method is useful for identifying
XX CC expression profiles which serve as useful diagnostic markers as well as
XX CC markers that can be used to monitor disease states, disease progression,
XX CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
XX CC for this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1348 BP; 370 A; 345 C; 341 G; 291 T; 0 U; 1 Other;
XX
XX
XX Query Match 29.5%; Score 34.2; DB 6; Length 1348;
XX Best Local Similarity 58.3%; Pred. No. 0.27;
XX Matches 60; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
XX
XX QY 10 ACTGCTGGAATACCTTTGTAGAAAACGAGAAAGACTTCAAGCCTGGAGAGGCTGC 69
XX Db 554 ACTGCTGGACACCTTTGTGACACACAGGATATCCCTTCAGCCCTGGAGTGAAGTAG 613
XX
XX QY 70 ATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCT 112
XX Db 614 ATGAGCACAGCCAGCCCTGAGTGGAGAGGCTGGCGGCGCATTTCT 656
XX
XX
XX RESULT 12
XX AAA12409
XX ID AAA12409 standard; cDNA, 1446 BP.
XX
XX AC AAA12409;
XX
XX DT 25-JUL-2000 (first entry)
XX
XX XX cDNA encoding a human RNA-associated protein.
XX
XX DE Human; RNA-associated protein; cell proliferation; cancer; inflammation;
XX KM immune response; reproductive disorder; actinic keratosis;
XX KM atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;
XX KM mixed connective tissue disease; myelofibrosis; primary thrombocytopenia;

KM paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;
KM trauma; ss.
XX
XX OS Homo sapiens.
XX
XX
XX Key location/Qualifiers
XX FT CDS 282..1446
XX FT /*tag= a
XX FT /product= "RNA-associated protein"
XX
XX PN MO200015799-A2.
XX
XX XX 23-MAR-2000.
XX
XX PD 17-SEP-1999; 99WO-US021688.
XX
XX PF 17-SEP-1998; 98US-00156039.
XX
XX PR 17-SEP-1998; 98US-00156039.
XX PR 22-SEP-1998; 98US-00158720.
XX PR 04-NOV-1998; 98US-0016815.
XX PR 08-APR-1999; 99US-0128660P.
XX
XX XX
XX PA (INCY-) INCYTE PHARM INC.
XX
XX XX Tang YT, Corley NC, Guejler KJ, Gorgone GA, Patterson C;
XX PI Hillman JI, Baughn MR, Lai P, Azimzal Y, Yue H, Yang J;
XX WPI; 2000-271437/23.
XX
XX DR F-PsDB; AAY84437.
XX
XX PT New polypeptides and polynucleotides, useful for preventing and treating
XX PT a disorder associated with increased or decreased expression of RNA
XX PT associated proteins.
XX
XX PS Claim 9; Page 119; 131bp; English.
XX
XX XX
XX CC The present sequence encodes a human RNA-associated protein. The
XX CC expression of RNA-associated proteins is closely associated with
XX CC reproductive tissues, nervous tissues, cell proliferation including
XX CC cancer, inflammation and immune responses, and so they may be used for
XX CC diagnosis, treatment or prevention of cell proliferative,
XX CC immune/inflammatory disorders, and reproductive disorders. Diseases and
XX CC disorders which may be treated include actinic keratosis,
XX CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
XX CC connective tissue disease, myelofibrosis, paroxysmal nocturnal
XX CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocytopenia
XX CC and cancers, and trauma
XX
XX SQ Sequence 1446 BP; 370 A; 374 C; 393 G; 308 T; 0 U; 1 Other;
XX
XX
XX Query Match 29.5%; Score 34.2; DB 3; Length 1446;
XX Best Local Similarity 58.3%; Pred. No. 0.28;
XX Matches 60; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
XX
XX QY 10 ACTGCTGGAATACCTTTGTAGAAAACGAGAAAGACTTCAAGCCTGGAGAGGCTGC 69
XX Db 1315 ACTGCTGGACACCTTTGTGACACACAGGATATCCCTTCAGCCCTGGATGACTAG 1374
XX
XX QY 70 ATGAAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCT 112
XX Db 1375 AGGAGCACAGCCAGCCCTGAGTGGAGAGGCTGGCGGCGCATTTCT 1417
XX
XX
XX RESULT 13
XX AA220856
XX ID AA220856 standard; cDNA, 1534 BP.
XX
XX AC AA220856;
XX
XX DT 09-DEC-1999 (first entry)
XX
XX XX Polynucleotide sequence of the Ip547_4 clone.
XX
XX DE Polynucleotide sequence of the Ip547_4 clone.
XX
XX KM secreted protein; cDNA library; clone; transmembrane protein;

KM signal sequence cloning; hybridization cloning; gene therapy; receptor;
KM ds.
XX Homo sapiens.
OS
FH Key 51.1205 Location/Qualifiers
FT CDS
FT sig_peptide /tag= a /product= "Ip547_4"
FT mat_peptide /tag= b /product= "882_1205"
FT /tag= c
XX WO9942470-A1.
XX 26-AUG-1999.
XX
XX 18-FEB-1999; 99WO-US003458.
XX
XX 18-FEB-1998; 98US-0075038P.
XX 17-FEB-1999; 99US-00251600.
XX
XX (GENY) GENETICS INST INC.
XX
XX Jacobs K, Mccoy JM, Lavallie BR, Collins-Racie LA, Merberg D;
PI Treacy M, Agostino MJ, Steininger RJ;
XX
XX WPI: 1999-518580/43.
XX P-PSDB; AA142383.
XX
XX New polynucleotides encoding human secreted proteins used for
PT therapeutic, diagnostic and research purposes.
XX
XX Claim 16; Page 104; 125pp; English.
XX
XX This is the polynucleotide sequence of the clone Ip547_4, which was
CC isolated from a human fetal brain cDNA library using methods which are
CC selective for cDNAs encoding secreted proteins, or by identification as a
CC secreted or transmembrane protein on the basis of computer analysis of
CC the amino acid sequence of the encoded protein. The PNs and proteins of
CC the invention are predicted to have biological activities which would
CC make them suitable for treating, preventing or ameliorating medical
CC conditions in humans and animals, although no supporting data is given.
CC Suggested activities include nutritional activity, cytokine and cell
CC proliferation/differentiation activity, immune stimulating (e.g. as
CC vaccines) or suppressing activity, hematopoiesis regulating activity,
CC tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, hemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumor
CC invasion suppressor activity, and tumor inhibition activity. The PNs are
CC also stated to be useful for gene therapy. Other activities include
CC inhibiting the growth, infection or function of bacteria, fungi, viruses
CC and other parasites; effecting bodily characteristics such as, e.g.
CC weight, color, skin, etc., effecting biorhythms or circadian cycles;
CC enhancing fertility; treatment of depression; treatment of pain; hormonal
CC or endocrine activity
XX
XX Sequence 1534 BP; 495 A; 363 C; 345 G; 331 T; 0 U; 0 Other;
SQ
Query Match 29.5%; Score 34.2; DB 2; Length 1534;
Best Local Similarity 58.3%; Pred. No. 0.29; Indels 0; Gaps 0;
Matches 60; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 10 ACTGCTGATATCTTTGTAGAAACACCAAGAACTTCAAGCTGGAAGGCTGC 69
DB 1084 ACTGCTGGAACACCTTTGTGACACCAAGGAGTGTCTTCCAGCCCTGGATGACTG 1143
QY 70 ATGAAATTCAGTTCCTCTCCAGACAGCTGGCGCATCT 112
DB 1144 ATGAGCACAGCCCAAGACCTGAGTGGGAGGCTGGCGGCATTTCT 1186

RESULT 14
AAS59293
ID AAS59293 standard; cDNA, 1534 BP.
XX
XX
AC AAS59293;
XX
XX 16-JAN-2002 (first entry)
DT
XX
XX
DE Human cDNA encoding a secreted protein Ip547_4.
XX
XX Human; secreted protein; ss; antiinflammatory; immunosuppressive;
KW neutropic; neuroprotective; antiarthritic; antimicrobial; vulnerary;
KW cytostatic; antidiabetic; virucide; antiinfectivity; anticonvulsant;
KW vasotropic; antiparkinsonian; immunostimulant; dermatological;
KW antirheumatic; antitumor; anticancer; osteopathic; tranquilizer;
KW cerebroprotective; cytokine; cell proliferation; cell differentiation;
KW immune deficiency; severe combined immunodeficiency; SCID; tumor;
KW autoimmune disorder; multiple sclerosis; rheumatoid arthritis;
KW graft-versus-host disease; myeloid deficiency; wound healing; ulcer;
KW periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; infection; cardiac disease;
KW stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
KW food supplement.
XX
XX Homo sapiens.
XX
XX WO200175066-A2.
XX
XX 11-OCT-2001.
XX
XX 22-MAR-2001; 2001WO-US009369.
XX
XX 30-MAR-2000; 2000US-00539330.
XX 04-DEC-2000; 2000US-00729674.
XX
XX (GENY) GENETICS INST INC.
XX
XX Jacobs K, Mccoy JM, Lavallie B, Collins-Racie LA, Evans C;
PI Treacy M, Agostino MJ, Steininger RJ, Spaulding V, Wong GG, Clark H;
PI Pectel K, Merberg D;
XX
XX WPI: 2001-639363/73.
XX P-PSDB; AAU39075.
XX
XX Secreted human proteins, useful as vaccine for treating various diseases
PT such as autoimmune disorders (e.g. multiple sclerosis), and nervous
PT system disorders (e.g. stroke).
XX
XX Disclosure; Page 580; 619pp; English.
XX
XX The invention relates to novel human secreted proteins, the nucleic acids
CC encoding them. The protein may exhibit cytokine, cell proliferation or
CC cell differentiation activity or may induce production of other cytokines
CC in certain cell populations and may exhibit immune stimulating or immune
CC suppressing activity, which is useful for the treatment of various immune
CC deficiencies and disorders e.g. severe combined immunodeficiency (SCID),
CC autoimmune disorders e.g. multiple sclerosis, systemic lupus
CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation.
CC The proteins are also useful in the treatment of diseases and disorders
CC including tissue, skin and organ transplantation and in graft-versus-host
CC diseases (GVHD), in the induction of tumour immunity, myeloid or lymphoid
CC cell deficiencies, wound healing and tissue repair, in the treatment of
CC burns, incisions and ulcers; as well as in treatment of periodontal
CC disease, osteoporosis or osteoarthritis, mediated by inflammatory
CC processes, diseases of the peripheral nervous system, Alzheimer's,
CC Parkinson's disease, Huntington's disease, amyotrophic lateral
CC sclerosis, and Shy-Drager syndrome, infections, infection of cardiac and
CC central nervous system vessel e.g. stroke, sepsis, inflammatory bowel
CC disease, ulcers, bone regeneration. The protein, having activin- or
CC inhibin-related activities is useful as a contraceptive based on the
CC ability of inhibin to decrease fertility in female mammals and decrease
CC spermatogenesis in male mammals. The proteins and nucleic acids are also
CC useful as food supplements. The present sequence encodes a secreted

CC Protein of the invention
 XX
 SQ Sequence 1534 BP; 495 A; 363 C; 345 G; 331 T; 0 U; 0 Other;
 Query Match 29.5%; Score 34.2; DB 4; Length 1534;
 Best Local Similarity 58.3%; Pred. No. 0.29;
 Matches 60; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
 QY 10 ACTGCTGGAATCTTTGTATGAAACCCGAAAGACTTTCAAGCCTGGAGGGCTGC 69
 DB 1084 ACTGCTGGACACCTTTGTGTGACCAACGAGATGTCCTTCAGCCTGGAGATGACTAG 1143
 QY 70 ATGAATTCAGTCTGCTCCAGACAGCCTTGGGCGCATCCT 112
 DB 1144 ATGAGCACAGCCAAAGACTTGAGTGAGGCTGCGGCGCATTCCT 1186
 RESULT 15
 ABA90962
 ID ABA90962 standard; cDNA; 1534 BP.
 AC ABA90962;
 AT 14-FEB-2002 (first entry)
 DE Human polynucleotide SEQ ID NO 173.
 XX
 KW Human; clone bd306-7; clone yb8-1; ATCC number 98599; gene therapy;
 KW immune disorder; bacterial infection; fungal infection; cancer; tumour;
 KW autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibit;
 KW osteoporosis; osteoarthritis; nervous system disorder; neuropathy;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;
 KW ischaemia-reperfusion injury; stroke; sepsis; arthritis; vulnery;
 KW haemophilia; cardiac infarction; inflammatory bowel disease; chemotactic;
 KW Crohn's disease; cytostatic; anti-inflammatory; immunomodulator;
 KW neuroprotective; haemostatic; thrombolytic; anti-inflammatory; ss.
 OS Homo sapiens.
 XX
 PN US2001039335-A1.
 XX
 PD 08-NOV-2001.
 XX
 PF 04-DEC-2000; 2000US-00729674.
 XX
 PR 26-NOV-1997; 97US-0126425P.
 PR 04-DEC-1997; 97US-0067454P.
 PR 20-DEC-1997; 97US-0068379P.
 PR 02-JAN-1998; 98US-0070346P.
 PR 07-JAN-1998; 98US-0070643P.
 PR 08-JAN-1998; 98US-0070755P.
 PR 13-JAN-1998; 98US-0071304P.
 PR 22-JAN-1998; 98US-0072134P.
 PR 30-JAN-1998; 98US-0073095P.
 PR 18-FEB-1998; 98US-0075038P.
 PR 23-NOV-1998; 98US-00197886.
 PR 30-MAR-2000; 2000US-00539330.
 XX
 PA (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LAVA/) LAVALLIE E R.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREB/) TREACY M.
 PA (AGOS/) AGOSTINO M J.
 PA (STEL/) STEININGER R J.
 PA (SPAU/) SPAULDING V.
 PA (WONG/) WONG G G.
 PA (CLAR/) CLARK H.
 PA (FECH/) FECHTEL K.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
 PI Wong GG, Clark H, Fechtel K;
 XX
 DR WPI: 2002-040725/05.
 XX P-PSDB: ABB55784.
 XX
 PT New secreted proteins and encoding polynucleotides, useful in gene
 PT therapies, particularly for preventing or treating autoimmune disorders,
 PT cancer, graft-versus-host disease, wound, osteoporosis, stroke or
 PT inflammations.
 XX
 PS Disclosure; Page 310-311; 349pp; English.
 XX
 CC The invention relates to isolated polynucleotides (ABA90876-ABA90968 and
 CC ABA90980) and encoded proteins (ABB55698-ABB55800), especially
 CC polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and
 CC proteins SEQ ID NO 2 (ABB55698) and SEQ ID NO 20 (ABB55707) contained in
 CC clones bd306-7 and yb8-1 respectively and the clones bd306-7 and yb8-1
 CC are deposited with the American Type Culture Collection (ATCC) with
 CC accession number 98599. The polynucleotides and encoded polypeptides have
 CC cytostatic, anti-inflammatory, immunomodulator, vulnery,
 CC neuroprotective, activin, inhibit, chemotactic, haemostatic, thrombolytic
 CC and anti-inflammatory activity and acting as cytokine modulators,
 CC haematopoiesis regulators, tissue growth modulators and/or cadexin
 CC suppressors. The polypeptides and polynucleotides are useful in gene
 CC therapies, particularly for preventing, treating or ameliorating any of
 CC the following diseases: immune deficiency and disorders; e.g. bacterial
 CC or fungal infections, autoimmune disorders, cancer, systemic lupus
 CC erythematosus or graft-versus-host disease, myeloid or lymphoid cell
 CC deficiencies; wound, burns, incisions and ulcers, osteoporosis or
 CC osteoarthritis; central and peripheral nervous system diseases and
 CC neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's
 CC disease, amyotrophic lateral sclerosis or Shy-Drager syndrome;
 CC haemophilia, cardiac infarction or stroke, inflammations, shock, sepsis
 CC or systemic inflammatory response syndrome, ischaemia-reperfusion injury,
 CC endotoxin lethality, arthritis, inflammatory bowel disease or Crohn's
 CC disease; or tumours or cancers, pemphigus vulgaris or pemphigus foliaceus
 XX
 SQ Sequence 1534 BP; 495 A; 363 C; 345 G; 331 T; 0 U; 0 Other;
 QY
 Query Match 29.5%; Score 34.2; DB 6; Length 1534;
 Best Local Similarity 58.3%; Pred. No. 0.29;
 Matches 60; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
 QY 10 ACTGCTGGAATCTTTGTATGAAACCCGAAAGACTTTCAAGCCTGGAGGGCTGC 69
 DB 1084 ACTGCTGGACACCTTTGTGTGACCAACGAGATGTCCTTCAGCCTGGAGATGACTAG 1143
 QY 70 ATGAATTCAGTCTGCTCCAGACAGCCTTGGGCGCATCCT 112
 DB 1144 ATGAGCACAGCCAAAGACTTGAGTGAGGCTGCGGCGCATTCCT 1186

Search completed: March 12, 2004, 18:08:10
 Job time : 47.4502 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 17:39:25 ; Search time 8.63095 Seconds
(without alignments)
7458.540 Million cell updates/sec

Title: US-09-966-880A-14

Perfect score: 116
Sequence: 1 attatcttaccgctgcgaat.....agcttcgcgcacatccttgc 116

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| C 1 | 27.2 | 23.4 | 2088 | 4 | US-09-513-057C-4 |
| C 2 | 27.2 | 23.4 | 2518 | 4 | US-09-513-057C-1 |
| C 3 | 27.2 | 23.4 | 2606 | 4 | US-09-334-827B-3 |
| C 4 | 27.2 | 23.4 | 4221 | 4 | US-09-513-057C-3 |
| C 5 | 27.2 | 23.4 | 4221 | 4 | US-09-513-057C-34 |
| C 6 | 27.2 | 23.4 | 8302 | 4 | US-09-334-827B-34 |
| 7 | 26.8 | 23.1 | 1125 | 1 | US-07-928-462-1 |
| 8 | 26.8 | 23.1 | 1125 | 3 | US-08-273-247-1 |
| 9 | 26.4 | 22.8 | 289 | 4 | US-09-313-294A-2945 |
| 10 | 26.2 | 22.6 | 80246 | 3 | US-09-078-294-4 |
| 11 | 26.2 | 22.6 | 80595 | 3 | US-09-078-294-3 |
| C 12 | 26.2 | 22.4 | 246 | 4 | US-09-543-681A-3465 |
| C 13 | 26.2 | 22.4 | 4944 | 1 | US-08-623-679-6 |
| C 14 | 26.2 | 22.4 | 4944 | 1 | US-08-623-679-6 |
| C 15 | 26.2 | 22.4 | 4944 | 3 | US-08-333-774-8 |
| C 16 | 26.2 | 22.4 | 4944 | 3 | US-09-181-030-6 |
| C 17 | 26.2 | 22.4 | 4944 | 4 | US-09-534-242-6 |
| C 18 | 26.2 | 22.4 | 4944 | 4 | US-09-164-854-6 |
| C 19 | 26.2 | 22.4 | 4944 | 4 | US-09-164-854-6 |
| C 20 | 26.2 | 22.4 | 5055 | 3 | US-08-623-679-8 |
| C 21 | 26.2 | 22.4 | 5055 | 3 | US-08-333-774-8 |
| C 22 | 26.2 | 22.4 | 5055 | 3 | US-09-181-030-8 |
| C 23 | 26.2 | 22.4 | 5055 | 4 | US-09-534-242-8 |
| C 24 | 26.2 | 22.4 | 5055 | 4 | US-09-164-854-8 |
| C 25 | 26.2 | 22.4 | 5746 | 2 | US-08-637-759B-38 |
| C 26 | 26.2 | 22.4 | 5746 | 3 | US-08-871-355A-38 |
| C 27 | 26.2 | 22.4 | 5746 | 4 | US-09-201-945-38 |

| | | | | | | |
|------|------|------|---------|---|---------------------|--------------------|
| C 28 | 25.4 | 21.9 | 452 | 1 | US-08-158-189-6 | Sequence 6, Appli |
| C 29 | 25.4 | 21.9 | 452 | 4 | US-09-016-434-1198 | Sequence 1198, Ap |
| C 30 | 25.4 | 21.9 | 627 | 4 | US-09-134-001C-1612 | Sequence 1612, Ap |
| C 31 | 25.4 | 21.9 | 1100 | 3 | US-09-248-335-53 | Sequence 53, Appl |
| C 32 | 25.2 | 21.7 | 508 | 4 | US-09-621-17979 | Sequence 17979, A |
| C 33 | 25.2 | 21.6 | 15472 | 4 | US-09-453-702B-71 | Sequence 71, Appl |
| C 34 | 25.2 | 21.6 | 1830121 | 4 | US-09-557-884-1 | Sequence 1, Appli |
| C 35 | 25.2 | 21.6 | 1830121 | 4 | US-09-643-990A-1 | Sequence 1, Appli |
| C 36 | 24.8 | 21.4 | 339 | 4 | US-09-621-12932 | Sequence 12932, A |
| C 37 | 24.8 | 21.4 | 684 | 4 | US-09-328-352-4049 | Sequence 4049, Ap |
| C 38 | 24.8 | 21.4 | 1001 | 4 | US-09-641-638-131 | Sequence 131, Appl |
| C 39 | 24.8 | 21.4 | 1318 | 4 | US-09-443-041A-25 | Sequence 25, Appl |
| C 40 | 24.8 | 21.4 | 5763 | 4 | US-08-956-171E-505 | Sequence 505, App |
| C 41 | 24.6 | 21.2 | 416 | 4 | US-09-621-976-18848 | Sequence 18848, A |
| C 42 | 24.6 | 21.2 | 567 | 4 | US-09-621-976-1240 | Sequence 1240, Ap |
| C 43 | 24.6 | 21.2 | 1664976 | 4 | US-08-916-421B-1 | Sequence 1, Appli |
| C 44 | 24.4 | 21.0 | 207 | 4 | US-09-134-001C-21 | Sequence 21, Appl |
| C 45 | 24.4 | 21.0 | 511 | 4 | US-09-621-976-3532 | Sequence 3532, Ap |

ALIGNMENTS

```
RESULT 1
US-09-513-057C-4/c
; Sequence 4, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:
; APPLICANT: Wagner, et al.
; TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
; FILE REFERENCE: 1505-54357
; CURRENT APPLICATION NUMBER: US/09/513,057C
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 2088
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-513-057C-4

Query Match      23.4%; Score 27.2; DB 4; Length 2088;
Best Local Similarity 52.7%; Pred. NO. 3.5;
Matches 59; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY      3 TATTTTACTGCTGGAATACCTTTGTGAAACCAAGAACTTCAAGCTGGGA 62
DB      184 TGTGTACTACGAGATTCATCTCTCCATGATCAACCACTCTGAGAGATCTAA 125

QY      63 GGGCTGATGAAATTCAGTTCGTCTCCAGACGCTTGGCGCATCTTT 114
DB      124 GCTGCTATAAAGAGCATCTGTTTCTAGAGAGAGCTTAGGCGCTCTTT 73

RESULT 2
US-09-513-057C-1/c
; Sequence 1, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:
; APPLICANT: Wagner, et al.
; TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
; FILE REFERENCE: 1505-54357
; CURRENT APPLICATION NUMBER: US/09/513,057C
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 2518
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (284)..(2371)
```

OTHER INFORMATION:
US-09-513-057C-1

Query Match 23.4%; Score 27.2; DB 4; Length 2518;
Best Local Similarity 52.7%; Pred. No. 3.8;
Matches 59; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 3 TATTTTACTGCTGGAATCTTTGTGAGAAAACCCAGAAAAGAACTTCAAGCCTGGGAA 62
DB 467 TGTGTTACTACGAGAAATTCATGCTTCATGATCACCAACCTCTGAGAAGAAATACTTA 408
QY 63 GGGCTGATGAAATTCAGTTCTCTCTCCAGACAGCTTCGGGCACTCTTT 114
DB 407 GCTGCTCATTAAGAGCCATCTTGTTCCTAGAGAGAGCTCTAGGCCCTCTTT 356

RESULT 3

US-09-234-827B-3/C
Sequence 3, Application US/09234827B
Patent No. 6448471
GENERAL INFORMATION:
APPLICANT: Puzio, Piotr S.
APPLICANT: Grunfelder, Florian M.W.
TITLE OF INVENTION: Nematode feeding structure specific gene and its
TITLE OF INVENTION: application to produce nematode resistant plants
FILE REFERENCE: U-012084-2
CURRENT APPLICATION NUMBER: US/09/234,827B
CURRENT FILING DATE: 1999-01-21
PRIOR APPLICATION NUMBER: US 60/072,142
PRIOR FILING DATE: 1998-01-22
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 2606
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: CDS
LOCATION: (328)..(2412)
US-09-234-827B-3

Query Match 23.4%; Score 27.2; DB 4; Length 2606;
Best Local Similarity 52.7%; Pred. No. 3.9;
Matches 59; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 3 TATTTTACTGCTGGAATCTTTGTGAGAAAACCCAGAAAAGAACTTCAAGCCTGGGAA 62
DB 511 TGTGTTACTACGAGAAATTCATGCTTCATGATCACCAACCTCTGAGAAGAAATACTTA 452
QY 63 GGGCTGATGAAATTCAGTTCTCTCTCCAGACAGCTTCGGGCACTCTTT 114
DB 451 GCTGCTCATTAAGAGCCATCTTGTTCCTAGAGAGAGCTCTAGGCCCTCTTT 400

RESULT 4

US-09-513-057C-3/C
Sequence 3, Application US/09513057C
Patent No. 6433251
GENERAL INFORMATION:
APPLICANT: Wagner, et al.
TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
FILE REFERENCE: 1505-54357
CURRENT APPLICATION NUMBER: US/09/513,057C
CURRENT FILING DATE: 2000-02-24
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentin version 3.1
SEQ ID NO 3
LENGTH: 4221
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(142)

OTHER INFORMATION:

NAME/KEY: 5'UTR
LOCATION: (143)..(425)
OTHER INFORMATION:
NAME/KEY: exon
LOCATION: (426)..(644)
OTHER INFORMATION:

NAME/KEY: exon
LOCATION: (1007)..(1803)
OTHER INFORMATION:
NAME/KEY: exon
LOCATION: (2984)..(3037)
OTHER INFORMATION:
NAME/KEY: exon
LOCATION: (3128)..(4142)
OTHER INFORMATION:
NAME/KEY: Intron
LOCATION: (645)..(1006)
OTHER INFORMATION:
NAME/KEY: Intron
LOCATION: (1804)..(2983)
OTHER INFORMATION:
NAME/KEY: Intron
LOCATION: (3038)..(3127)
OTHER INFORMATION:
NAME/KEY: 3'UTR
LOCATION: (4146)..(4221)
OTHER INFORMATION:
US-09-513-057C-3

Query Match 23.4%; Score 27.2; DB 4; Length 4221;
Best Local Similarity 52.7%; Pred. No. 4.8;
Matches 59; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 3 TATTTTACTGCTGGAATCTTTGTGAGAAAACCCAGAAAAGAACTTCAAGCCTGGGAA 62
DB 609 TGTGTTACTACGAGAAATTCATGCTTCATGATCACCAACCTCTGAGAAGAAATACTTA 550
QY 63 GGGCTGATGAAATTCAGTTCTCTCTCCAGACAGCTTCGGGCACTCTTT 114
DB 549 GCTGCTCATTAAGAGCCATCTTGTTCCTAGAGAGAGCTCTAGGCCCTCTTT 498

RESULT 5

US-09-513-057C-34/C
Sequence 34, Application US/09513057C
Patent No. 6433251
GENERAL INFORMATION:
APPLICANT: Wagner, et al.
TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
FILE REFERENCE: 1505-54357
CURRENT APPLICATION NUMBER: US/09/513,057C
CURRENT FILING DATE: 2000-02-24
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentin version 3.1
SEQ ID NO 34
LENGTH: 4221
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)..(425)
OTHER INFORMATION:
NAME/KEY: exon
LOCATION: (426)..(644)
OTHER INFORMATION:
NAME/KEY: Intron
LOCATION: (645)..(1006)
OTHER INFORMATION:
NAME/KEY: exon
LOCATION: (1007)..(1803)
OTHER INFORMATION:
NAME/KEY: Intron

LOCATION: (1804)..(2983)
OTHER INFORMATION:
NAME/KEY: exon
LOCATION: (2984)..(3035)
OTHER INFORMATION:
NAME/KEY: Intron
LOCATION: (3036)..(3125)
OTHER INFORMATION:
NAME/KEY: exon
LOCATION: (3126)..(4145)
OTHER INFORMATION:
NAME/KEY: 3'UTR
LOCATION: (4146)..(4221)
OTHER INFORMATION:
US-09-513-057C-34

Query Match 23.4%; Score 27.2; DB 4; Length 4221;
Best Local Similarity 52.7%; Pred. No. 4.8;
Matches 59; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 3 TATTTTACTGCTGGAATCTTTGTAGAAACACGAAAGAACTTCAAGCTGGGA 62
DB 609 TGTGTACTACGAGAAATCTCTTCATGATCACCACCTTGAAGGAACTCTAA 550
QY 63 GGGCTGATGAAATTCAGTTCCTCTCCAGACGCTGGCGCATCTTT 114
DB 549 GCTGCTATTAAGAGCATCTTGTCTTAGAGAGAGCTTAGGCCCTCTTT 498

RESULT 6
US-09-234-827B-1/c
Sequence 1, Application US/09234827B
Patent No. 6448471
GENERAL INFORMATION:
APPLICANT: Puzio, Piotr S.
APPLICANT: Grunfelder, Florian M.W.
TITLE OF INVENTION: Nematode feeding structure specific gene and its
FILE REFERENCE: U-012084-2
CURRENT APPLICATION NUMBER: US/09/234,827B
CURRENT FILING DATE: 1999-01-21
PRIOR APPLICATION NUMBER: US 60/072,142
PRIOR FILING DATE: 1998-01-22
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 8302
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: CDS
LOCATION: (4038)..(4256)
NAME/KEY: CDS
LOCATION: (4807)..(5604)
NAME/KEY: CDS
LOCATION: (6777)..(6827)
NAME/KEY: CDS
LOCATION: (6919)..(7935)
US-09-234-827B-1

Query Match 23.4%; Score 27.2; DB 4; Length 8302;
Best Local Similarity 52.7%; Pred. No. 6.5;
Matches 59; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 3 TATTTTACTGCTGGAATCTTTGTAGAAACACGAAAGAACTTCAAGCTGGGA 62
DB 4221 TGTGTACTACGAGAAATCTCTTCATGATCACCACCTTGAAGGAACTCTAA 4162
QY 63 GGGCTGATGAAATTCAGTTCCTCTCCAGACGCTGGCGCATCTTT 114
DB 4161 GCTGCTATTAAGAGCATCTTGTCTTAGAGAGAGCTTAGGCCCTCTTT 4110

RESULT 7

US-07-928-462-1
Sequence 1, Application US/07928462
Patent No. 5328996

GENERAL INFORMATION:

APPLICANT: Boyle, Michael D.P.
APPLICANT: Lottenberg, Richard
APPLICANT: Broder, Christopher C.
APPLICANT: von Merling, Gregory O.
TITLE OF INVENTION: Bacterial Plasmin Receptors as
TITLE OF INVENTION: Fibrinolytic Agents
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanichik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/928,462
FILING DATE: 19920810
CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/524,411
FILING DATE: 16-MAY-1990
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/330,849

FILING DATE: 29-MAR-1989

ATTORNEY/AGENT INFORMATION:

NAME: Saliwanichik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF/S&S-13.C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1125 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: both

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

ORGANISM: Streptococcus pyogenes

STRAIN: M untypable

INDIVIDUAL ISOLATE: 64/14

IMMEDIATE SOURCE:

CLONE: pRL015

FEATURE:

NAME/KEY: CDS
LOCATION: 115..1122

FEATURE:

NAME/KEY: mat_peptide
LOCATION: 115..1122

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /codon start= 115

OTHER INFORMATION: /function= "High-affinity binding of

OTHER INFORMATION: /product= "Streptococcal plasmin receptor"

OTHER INFORMATION: /evidence= EXPERIMENTAL

OTHER INFORMATION: /gene= "plr"

OTHER INFORMATION: /number= 1

OTHER INFORMATION: /label= plr

US-07-928-462-1

Query Match 23.1%; Score 26.8; DB 1; Length 1125;
Best Local Similarity 59.0%; Pred. No. 3.7;
Matches 46; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 35 CCAGGAAAGACTTCAAGCCCTGGAGAGGCTGCATGAATTCAGTTCCTCCAG 94
DB 899 CTATGAAAGCTGCTTCAACGATAGCTTCGTTACACTGAAGATCCATCGTTCTTCAG 958

QY 95 ACAGCTTCGGCGCATCCT 112
DB 959 ATATCGTAGCGGTATCAT 976

US-08-273-247-1
; Sequence 1, Application US/08273247
; Patent No. 6136323
; GENERAL INFORMATION:
; APPLICANT: Boyle, Michael D.P.
; APPLICANT: Lotenberg, Richard
; APPLICANT: Broder, Christopher C.
; APPLICANT: von Mering, Gregory O.
; TITLE OF INVENTION: Bacterial Plasmin Receptors as
; TITLE OF INVENTION: Fibrinolytic Agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,247
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/524,411
; FILING DATE: 16-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/330,849
; FILING DATE: 29-MAR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF/S&S-13.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1125 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
; STRAIN: M untypable
; INDIVIDUAL ISOLATE: 64/14
; IMMEDIATE SOURCE:
; CLONE: pRL015
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 115..1122
; FEATURE:

NAME/KEY: mac peptide
; LOCATION: 115..1122
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /codon_start=115
; OTHER INFORMATION: /function="High-affinity binding of
; OTHER INFORMATION: /plasma(ogen)"
; OTHER INFORMATION: /product="Streptococcal plasmin receptor"
; OTHER INFORMATION: /evidence=EXPERIMENTAL
; OTHER INFORMATION: /gene="plr"
; OTHER INFORMATION: /number=1
; OTHER INFORMATION: /label=plr

US-08-273-247-1
Query Match 23.1%; Score 26.8; DB 3; Length 1125;
Best Local Similarity 59.0%; Pred. No. 3.7;
Matches 46; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 35 CCAGGAAAGACTTCAAGCCCTGGAGAGGCTGCATGAATTCAGTTCCTCCAG 94
DB 899 CTATGAAAGCTGCTTCAACGATAGCTTCGTTACACTGAAGATCCATCGTTCTTCAG 958

QY 95 ACAGCTTCGGCGCATCCT 112
DB 959 ATATCGTAGCGGTATCAT 976

US-09-313-294A-2945
; Sequence 2945, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 2945
; LENGTH: 289
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: incyte ID No. 6476212 700282045H1
; NAME/KEY: unsure
; LOCATION: 22, 89
; OTHER INFORMATION: a, c, g, or other
; US-09-313-294A-2945

Query Match 22.8%; Score 26.4; DB 4; Length 289;
Best Local Similarity 52.9%; Pred. No. 2.9;
Matches 54; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 7 TTACTGTGATGATCTTTGTAGAAAACGAGAAAGAACTTTCAGGCTGGAGAGGC 66
DB 6 TATATCCATGATATACGCGGTAGAGAAAGAAAGATGATAGCTCGAAGCAGGGT 65

QY 67 TGCAAGAAATTCAGTTCCTCCGAGAGAGCTGGCGCC 108
DB 66 AGAGACATCAATTAAGGCTTCACAGCAGAGCTTTGCTCA 107

US-09-078-294-4
; Sequence 4, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.

;; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
;; FILE REFERENCE: Davies Col
;; CURRENT APPLICATION NUMBER: US/09/078,294
;; CURRENT FILING DATE: 1998-05-13
;; NUMBER OF SEQ ID NOS: 29
;; SOFTWARE: Patentin Ver. 2.0
;; SEQ ID NO 4
;; LENGTH: 80246
;; TYPE: DNA
;; ORGANISM: Nucleotide sequence of NC-contig
US-09-078-294-4

Query Match 22.6%; Score 26.2; DB 3; Length 80246;
Best Local Similarity 52.3%; Pred. No. 41;
Matches 58; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 5 TTTTCTGCTGGAATCTTTGTAGAAACGAAAGACTTTGAAGCCTGGGAAG 64
DB 67330 TTCTTCTCCAGTAGCAATATGCTTAAGGAAAGAAAGAGGGAAGC 67389
QY 65 GCTGATGAATTCAGTTCCTCCAGACAGCTTCGGCGCATCTTT 115
DB 67390 ACTGAACCAATTTGCTTTTGTACTGCTCAAGCTCAATGCAAGTTCT 67440

RESULT 11
US-09-078-294-3
; Sequence 3, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 80595
; TYPE: DNA
; ORGANISM: Nucleotide sequence of HC-contig
US-09-078-294-3

Query Match 22.6%; Score 26.2; DB 3; Length 80595;
Best Local Similarity 52.3%; Pred. No. 41;
Matches 58; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 5 TTTTCTGCTGGAATCTTTGTAGAAACGAAAGACTTTGAAGCCTGGGAAG 64
DB 67588 TTCTTCTCCAGTAGCAATATGCTTAAGGAAAGAAAGAGGGAAGC 67647
QY 65 GCTGATGAATTCAGTTCCTCCAGACAGCTTCGGCGCATCTTT 115
DB 67648 ACTGAACCAATTTGCTTTTGTACTGCTCAAGCTCAATGCAAGTTCT 67698

RESULT 12
US-09-543-681A-3465/C
; Sequence 3465, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709, 1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3465

;; LENGTH: 246
;; TYPE: DNA
;; ORGANISM: Proteus mirabilis
US-09-543-681A-3465

Query Match 22.4%; Score 26; DB 4; Length 246;
Best Local Similarity 62.1%; Pred. No. 3.8;
Matches 41; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 31 AAACACGAAAGACTTCAAGCTGGGAAGGCTGCATGAATTCAGTTCCT 90
DB 239 AAGAAATCAATTAATTAATGACCGAGTTAAATGTGAATTAATCACTGCTTCA 180
QY 91 CCAGAC 96
DB 179 CTGGAC 174

RESULT 13
US-08-623-679-6
; Sequence 6, Application US/08623679
; Patent No. 5674739
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: DIAGNOSIS, PREVENTION AND TREATMENT OF TUMOR
; TITLE OF INVENTION: PROGRESSION
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/623,679
; FILING DATE: 29-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/412,431
; FILING DATE: 29-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4944 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 346..4837
US-08-623-679-6

Query Match 22.4%; Score 26; DB 1; Length 4944;
Best Local Similarity 54.1%; Pred. No. 14;
Matches 53; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 ATTATTTTCTGCTGGAATCTTTGTAGAAACGAAAGACTTTGAAGCCTGG 60

Db 2441 ATGATGATTCTCGTATCAACATCCAGGAAACGAGATGCGCAAGAAAAGAGAGG 2500
 QY 61 AAGGCTGCATGAAATTCAGTTCGTCTCTCCAGACG 98
 Db 2501 AAAATACGATGCAAAATGCAAGATGCTGCTCAAGAAAG 2538

RESULT 14
 US-08-933-774-6
 ; Sequence 6, Application US/08933774A
 ; Patent No. 6025137
 ; GENERAL INFORMATION:

; APPLICANT: Shyjan, Andrew W.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS, PREVENTION
 ; TITLE OF INVENTION: AND TREATMENT OF TUMOR PROGRESSION
 ; FILE REFERENCE: 07334/004003
 ; CURRENT APPLICATION NUMBER: US/08/933,774A
 ; EARLIER FILING DATE: 1997-09-19
 ; EARLIER APPLICATION NUMBER: US 08/623,679
 ; EARLIER FILING DATE: 1996-03-29
 ; EARLIER APPLICATION NUMBER: US 08/412,431
 ; EARLIER FILING DATE: 1995-03-29
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 6
 ; LENGTH: 4944
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (346)...(4836)
 US-08-933-774-6

Query Match 22.4%; Score 26; DB 3; Length 4944;
 Best Local Similarity 54.1%; Pred. No. 14;
 Matches 53; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAATCTTTGTAGAAAACCGAAGAACTTCAAGCCTGGG 60
 Db 2441 ATGATGATTCTCGTATCAACATCCAGGAAACGAGATGCGCAAGAAAAGAGAGG 2500
 QY 61 AAGGCTGCATGAAATTCAGTTCGTCTCTCCAGACG 98
 Db 2501 AAAATACGATGCAAAATGCAAGATGCTGCTCAAGAAAG 2538

RESULT 15
 US-09-181-030-6
 ; Sequence 6, Application US/09181030
 ; Patent No. 6251597
 ; GENERAL INFORMATION:

; APPLICANT: Shyjan, Andrew W.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS, PREVENTION
 ; TITLE OF INVENTION: AND TREATMENT OF TUMOR PROGRESSION
 ; FILE REFERENCE: 07334/004005
 ; CURRENT APPLICATION NUMBER: US/09/181,030
 ; EARLIER FILING DATE: 1998-10-27
 ; EARLIER APPLICATION NUMBER: US 08/862,442
 ; EARLIER FILING DATE: 1997-05-23
 ; EARLIER APPLICATION NUMBER: US 08/623,679
 ; EARLIER FILING DATE: 1996-03-29
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 6
 ; LENGTH: 4944
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (346)...(4836)
 US-09-181-030-6

Query Match 22.4%; Score 26; DB 3; Length 4944;

Best Local Similarity 54.1%; Pred. No. 14;
 Matches 53; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
 QY 1 ATTATTTTACTGCTGGAATCTTTGTAGAAAACCGAAGAACTTCAAGCCTGGG 60
 Db 2441 ATGATGATTCTCGTATCAACATCCAGGAAACGAGATGCGCAAGAAAAGAGAGG 2500
 QY 61 AAGGCTGCATGAAATTCAGTTCGTCTCTCCAGACG 98
 Db 2501 AAAATACGATGCAAAATGCAAGATGCTGCTCAAGAAAG 2538

Search completed: March 13, 2004, 00:51:23
 Job time : 11.631 secs


```
RESULT 2
US-09-966-880A-10
; Sequence 10, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 6564
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-966-880A-10

Query Match          100.0%; Score 116; DB 9; Length 6564;
Best Local Similarity 100.0%; Pred. No. 1.1e-30;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATTATTTTACTGCTGGAATCTTTGTAGAAACACGAAAGAACTTCAAGCTGGG 60
DB      3156 ATTATTTTACTGCTGGAATCTTTGTAGAAACACGAAAGAACTTCAAGCTGGG 3215

QY      61 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTGGGCAATCCCTTTG 116
DB      3216 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTGGGCAATCCCTTTG 3271

RESULT 3
US-09-966-880A-35
; Sequence 35, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 11204
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-966-880A-35

Query Match          100.0%; Score 116; DB 9; Length 11204;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ATTATTTTACTGCTGGAATCTTTGTAGAAACACGAAAGAACTTCAAGCTGGG 60
DB      8371 ATTATTTTACTGCTGGAATCTTTGTAGAAACACGAAAGAACTTCAAGCTGGG 8430

QY      61 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTGGGCAATCCCTTTG 116
DB      8431 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTGGGCAATCCCTTTG 8486

RESULT 4
US-09-966-880A-7
; Sequence 7, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2818
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (80)...(673)
; NAME/KEY: 5'UTR
; LOCATION: (1)...(79)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (677)...(2818)
US-09-966-880A-7

Query Match          98.6%; Score 114.4; DB 9; Length 2818;
Best Local Similarity 99.1%; Pred. No. 3e-30;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ATTATTTTACTGCTGGAATCTTTGTAGAAACACGAAAGAACTTCAAGCTGGG 60
DB      507 ATTATTTTACTGCTGGAATCTTTGTAGAAACACGAAAGAACTTCAAGCTGGG 566

QY      61 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTGGGCAATCCCTTTG 116
DB      567 AAGGCTGCATGAAATTCAGTTGCTCTCCAGACAGCTTGGGCAATCCCTTTG 622

RESULT 5
US-09-966-880A-1
; Sequence 1, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
```

```

; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2440
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (93)...(686)
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)...(92)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (690)...(2440)
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(2440)
; OTHER INFORMATION: n = A,T,C or G
US-09-966-880A-1

Query Match      83.4%; Score 96.8; DB 9; Length 2440;
Best Local Similarity 89.7%; Pred. No. 5.5e-24;
Matches 104; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAATCTTTTGTAAGAAACCAAGAACTTCAAGCTGGG 60
DB 520 ACTATTTTACTGCTGGAATCTTTTGTAAGAACTTCAAGCTGGG 579
DB 580 AAGGCTGATGAATTCAGTTCTCTCCAGAGAGTTGGGCACTCTTTG 116
DB 580 AAGGCTGATGAATTCAGTTCTCTCTCCAGAGAGTTGGGCACTCTTTG 635

RESULT 6
US-09-796-692-9051
; Sequence 9051, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
```

```

; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9051
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-9051

Query Match      29.5%; Score 34.2; DB 9; Length 390;
Best Local Similarity 58.3%; Pred. No. 0.062;
Matches 60; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 10 ACTGCTGAATCTTTTGTAAGAAACCAAGAACTTCAAGCTGGGAAGGCTGC 69
DB 165 ACTGCTGGAACCTTTGTGACACACAGGATGTCCCTTCCAGCCCTGGATGACTAG 224
QY 70 ATGAAATTCAGTTCTCTCCAGACAGCTTGGCGCATCT 112
DB 225 ATGAGCACAGCCAAAGCTGATGAGGAGCTGGCGCATCT 267

RESULT 7
US-10-040-862-9051
; Sequence 9051, Application US/10040862
; Publication No. US2003078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Reltzer, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9051
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-9051

Query Match      29.5%; Score 34.2; DB 14; Length 390;
```

Best Local Similarity 58.3%; Pred. No. 0.062;
Matches 60; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 10 ACTGCTGGAATACCTTTGTAGAAAACCAAGAACTTTCAAGCTGGGAAGGCTGC 69
Db 165 ACTGCTGGGACACCTTTGTAGACCAACGAGGATGTCCCTTCAGCCCTGGATGACTAG 224

QY 70 ATGAAATTCAGTCTCTCTCCAGACAGCTTGGCGCATCT 112
Db 225 ATGAGCAGACCAAGACCTGAGTGGAGGCTGGCGCATCT 267

RESULT 8

US-10-057-475B-9051
; Sequence 9051, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannon, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aljun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Diane
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9051
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-9051

Query Match 29.5%; Score 34.2; DB 15; Length 390;
Best Local Similarity 58.3%; Pred. No. 0.062;
Matches 60; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 10 ACTGCTGGAATACCTTTGTAGAAAACCAAGAACTTTCAAGCTGGGAAGGCTGC 69
Db 165 ACTGCTGGGACACCTTTGTAGACCAACGAGGATGTCCCTTCAGCCCTGGATGACTAG 224

QY 70 ATGAAATTCAGTCTCTCTCCAGACAGCTTGGCGCATCT 112
Db 225 ATGAGCAGACCAAGACCTGAGTGGAGGCTGGCGCATCT 267

RESULT 9
US-10-154-884B-9051
; Sequence 9051, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannon, Jane
; APPLICANT: Reiter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9051
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-9051

Query Match 29.5%; Score 34.2; DB 15; Length 390;
Best Local Similarity 58.3%; Pred. No. 0.062;
Matches 60; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 10 ACTGCTGGAATACCTTTGTAGAAAACCAAGAACTTTCAAGCTGGGAAGGCTGC 69
Db 165 ACTGCTGGGACACCTTTGTAGACCAACGAGGATGTCCCTTCAGCCCTGGATGACTAG 224

QY 70 ATGAAATTCAGTCTCTCTCCAGACAGCTTGGCGCATCT 112
Db 225 ATGAGCAGACCAAGACCTGAGTGGAGGCTGGCGCATCT 267

RESULT 10
US-10-460-923-1
; Sequence 1, Application US/10460923
; Publication No. US20040009951A1
; GENERAL INFORMATION:
; APPLICANT: Malim, Michael H.
; APPLICANT: Sheehy, Ann M.
; APPLICANT: Harris, Reuben S.
; APPLICANT: Bishop, Kate N.
; APPLICANT: Neuberger, Michael S.
; APPLICANT: Gaddis, Nathan C.
; APPLICANT: Simon, James H.M.
; TITLE OF INVENTION: DNA Dendrimation Mediates Innate Immunity to Retroviral Infection
; FILE REFERENCE: 22253-74380
; CURRENT APPLICATION NUMBER: US/10/460,923
; CURRENT FILING DATE: 2003-06-13

```
; PRIOR APPLICATION NUMBER: US 60/388,513
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 60/472,952
; PRIOR FILING DATE: 2003-05-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-460-923-1

Query Match          29.5%; Score 34.2; DB 15; Length 1155;
Best Local Similarity 58.3%; Pred. No. 0.09; Indels 0; Gaps 0;
Matches 60; Conservative 0; Mismatches 43;

QY 10 ACTGCTGGAATCTTTGTAGAAAACCAAGAACTTCAAGCCTGGAGGGCTGC 69
DB 1034 ACTGCTGGGACACCTTTGTGACACCAAGGATGTCCCTTCAGCCCTGGAGTGAAGT 1093
QY 70 ATGAAATTCAGTTCCTCTCCAGACAGCTTCGGGCGCATCT 112
DB 1094 ATGAGCAGCAGCAAGACCTAGTGGGAGGCTGGGGCGCATCT 1136

RESULT 11
US-09-880-107-3282
; Sequence 3282, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scheer, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-MO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3282
; LENGTH: 1348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U03891
; NAME/KEY: unsure
; LOCATION: (1)-(1348)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-3282

Query Match          29.5%; Score 34.2; DB 9; Length 1348;
Best Local Similarity 58.3%; Pred. No. 0.11; Indels 0; Gaps 0;
Matches 60; Conservative 0; Mismatches 43;

QY 10 ACTGCTGGAATCTTTGTAGAAAACCAAGAACTTCAAGCCTGGAGGGCTGC 69
DB 554 ACTGCTGGGACACCTTTGTGACACCAAGGATGTCCCTTCAGCCCTGGAGTGAAGT 613
QY 70 ATGAAATTCAGTTCCTCTCCAGACAGCTTCGGGCGCATCT 112
DB 614 ATGAGCAGCAGCAAGCCCTAGTGGGAGGCTGGGGCGCATCT 656

RESULT 12
US-09-729-674-173
; Sequence 173, Application US/09729674
; Patent No. US2001003935A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavalie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Metberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Reckel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 173
; LENGTH: 1534
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-729-674-173

Query Match          29.5%; Score 34.2; DB 9; Length 1534;
Best Local Similarity 58.3%; Pred. No. 0.11; Indels 0; Gaps 0;
Matches 60; Conservative 0; Mismatches 43;

QY 10 ACTGCTGGAATCTTTGTAGAAAACCAAGAACTTCAAGCCTGGAGGGCTGC 69
DB 1084 ACTGCTGGGACACCTTTGTGACACCAAGGATGTCCCTTCAGCCCTGGAGTGAAGT 1143
QY 70 ATGAAATTCAGTTCCTCTCCAGACAGCTTCGGGCGCATCT 112
DB 1144 ATGAGCAGCAGCAAGACCTAGTGGGAGGCTGGGGCGCATCT 1186

RESULT 13
US-10-240-425-1099
; Sequence 1099, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Metzger, Jon C.
; APPLICANT: Scheer, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; PRIOR FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1099
; LENGTH: 201143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 AL022318
US-10-240-425-1099

Query Match          29.5%; Score 34.2; DB 12; Length 201143;
Best Local Similarity 58.3%; Pred. No. 0.93;
```


FEATURES
Source
Location/Qualifiers
1. .693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4336722"
/issue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 98.6%; Score 114.4; DB 10; Length 693;
Best Local Similarity 99.1%; Pred. No. 1.5e-26;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAATCTTTGTAGAAAACGAGAAAGCTTCAAGCCTGGG 60
DB 204 ATTATTTTACTGCTGGAATCTTTGTAGAAAACGAGAAAGCTTCAAGCCTGGG 263

QY 61 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCCTTTTG 116
DB 264 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCCTTTTG 319

RESULT 2
Bg686133 743 bp mRNA linear EST 01-MAY-2001
LOCUS Bg686133
DEFINITION 602638412F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4766234 5',
mRNA sequence.
ACCESSION Bg686133.1 GI:13917530
VERSION Bg686133.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM626 row: 9 column: 03
High quality sequence stop: 740.
Location/Qualifiers
1. .743
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4766234"
/issue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 98.6%; Score 114.4; DB 12; Length 743;
Best Local Similarity 99.1%; Pred. No. 1.5e-26;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAATCTTTGTAGAAAACGAGAAAGCTTCAAGCCTGGG 60
DB 505 ATTATTTTACTGCTGGAATCTTTGTAGAAAACGAGAAAGCTTCAAGCCTGGG 564

QY 61 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCCTTTTG 116
DB 565 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCCTTTTG 620

RESULT 3
Bg757089 820 bp mRNA linear EST 15-MAY-2001
LOCUS Bg757089
DEFINITION 602715124F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4855517 5',
mRNA sequence.
ACCESSION Bg757089
VERSION Bg757089.1 GI:14067742
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM704 row: 0 column: 06
High quality sequence stop: 675.
Location/Qualifiers
1. .820
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4855517"
/issue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 98.6%; Score 114.4; DB 12; Length 820;
Best Local Similarity 99.1%; Pred. No. 1.5e-26;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAATCTTTGTAGAAAACGAGAAAGCTTCAAGCCTGGG 60

Db 512 ATTATTTTACTGCTGGAATACCTTTGTAGAAAACCATGAAGAACTTCAAGCTGGG 571

QY 61 AAGGGCTGCATGAATAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 116

Db 572 AAGGGCTGCATGAATAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 627

RESULT 4

LOCUS BG758510

DEFINITION 872 bp mRNA linear EST 15-MAY-2001

602712721.F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853069.5', mRNA sequence.

ACCESSION BG758510

VERSION BG758510

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE NIH-MGC National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: L1CM2108 row: 1 column: 06
High quality sequence stop: 836.

FEATURES

source

Location/Qualifiers

1..872

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4853069"

/tissue_type="primary B-cells from tonsils (cell line)"

/lab_host="DH10B (phage-resistant)"

/clone_1lb="NIH_MGC_48"

/note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 98.6%; Score 114.4; DB 12; Length 872;

Best Local Similarity 99.1%; Pred. No. 1.5e-26;

Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAATACCTTTGTAGAAAACCATGAAGAACTTCAAGCTGGG 60

Db 493 ATTATTTTACTGCTGGAATACCTTTGTAGAAAACCATGAAGAACTTCAAGCTGGG 552

QY 61 AAGGGCTGCATGAATAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 116

Db 553 AAGGGCTGCATGAATAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 608

RESULT 5

LOCUS BQ065440

DEFINITION 953 bp mRNA linear EST 02-APR-2002

60065440

AGENCOURT 6855061 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5929977

5', mRNA sequence.

ACCESSION BQ065440

VERSION BQ065440.1 GI:19894486

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE NIH-MGC National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: L1CM2108 row: p column: 10
High quality sequence stop: 634.

FEATURES

source

Location/Qualifiers

1..953

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5929977"

/tissue_type="lymphoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_1lb="NIH_MGC_99"

/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 98.6%; Score 114.4; DB 13; Length 953;

Best Local Similarity 99.1%; Pred. No. 1.6e-26;

Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAATACCTTTGTAGAAAACCATGAAGAACTTCAAGCTGGG 60

Db 491 ATTATTTTACTGCTGGAATACCTTTGTAGAAAACCATGAAGAACTTCAAGCTGGG 550

QY 61 AAGGGCTGCATGAATAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 116

Db 551 AAGGGCTGCATGAATAATTCAGTTGCTCTCCAGACAGCTTCGGCGCATCTTTTG 606

RESULT 6

LOCUS BQ055935

DEFINITION 1052 bp mRNA linear EST 29-MAR-2002

60055935

AGENCOURT 6796291 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5808181

5', mRNA sequence.

ACCESSION BQ055935

VERSION BQ055935.1 GI:19815262

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE NIH-MGC National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt

CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LNCM2051 row: m column: 14
 High quality sequence stop: 665.
 Location/Qualifiers
 1..1052

FEATURES

source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5808181"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_93"
/note="Organ: lymph; Vector: pOTB7, Site_1: XhoI, Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

ORIGIN

Query Match 98.6%; Score 114.4; DB 13; Length 1052;
 Best Local Similarity 99.1%; Pred. No. 1.6e-26;
 Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAATCTTTGTAGAAAACCGAAAGACTTTCAAGCCTGGG 60
 |||||
 DB 491 ATTATTTTACTGCTGGAATCTTTGTAGAAAACCGAAAGACTTTCAAGCCTGGG 550
 |||||

QY 61 AAGGCGTCAGTAATAATTCAGTTCCTCCAGACAGCTTCGGCGCATCCTTTTG 116
 |||||
 DB 551 AAGGCGTCAGTAATAATTCAGTTCCTCCAGACAGCTTCGGCGCATCCTTTTG 606
 |||||

RESULT 7 1201 bp mRNA linear EST 31-MAY-2003
 AL559877
 LOCUS
 DEFINITION AL559877 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens CDNA
 clone CSODG003YB14 5-PRIME, mRNA sequence.
 ACCESSION AL559877
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1201)
 Li M.B., Gruber C., Jesse J. and Polayes D.
 Full-length CDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:12905793.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 104 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 6672.r for
 more information about this cluster, see
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODG003DA07QPL&cluster=6672.r)
 cgi-bin/cluster.cgi?seq=CSODG003DA07QPL&cluster=6672.r. Contact :
 Feng Liang Email: fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>
 Faraday Avenue Genoscope sequence ID : CSODG003DA07QPL.
 Location/Qualifiers

FEATURES

source

```

1..1201
/organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODG003YB14"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand CDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand CDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

ORIGIN

Query Match 98.6%; Score 114.4; DB 9; Length 1201;
 Best Local Similarity 99.1%; Pred. No. 1.6e-26;
 Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAATCTTTGTAGAAAACCGAAAGACTTTCAAGCCTGGG 60
 |||||
 DB 519 ATTATTTTACTGCTGGAATCTTTGTAGAAAACCGAAAGACTTTCAAGCCTGGG 578
 |||||

QY 61 AAGGCGTCAGTAATAATTCAGTTCCTCCAGACAGCTTCGGCGCATCCTTTTG 116
 |||||
 DB 579 AAGGCGTCAGTAATAATTCAGTTCCTCCAGACAGCTTCGGCGCATCCTTTTG 634
 |||||

RESULT 8

BG757392 693 bp mRNA linear EST 15-MAY-2001
 LOCUS
 DEFINITION 602711022P1 NIH_MGC_48 Homo sapiens CDNA clone IMAGE:4851580 5',
 mRNA sequence.
 ACCESSION BG757392
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 693)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LNCM1694 row: x column: 05
 High quality sequence stop: 693.
 Location/Qualifiers

FEATURES

source

```

1..693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4851580"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7, Site_1: XhoI,
Site_2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGCGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

ORIGIN

Query Match 97.2%; Score 112.8; DB 12; Length 693;
 Best Local Similarity 98.3%; Pred. No. 5.1e-26;
 Matches 114; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAATACCTTTGTAGAAAAACCAAGAAAGACTTTCAAGCCTGGG 60
 DB 509 ATTATTTTACTGCTGGAATACCTTTGTAGAAAAACCAAGAAAGACTTTCAAGCCTGGG 568

QY 61 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCCTTTTG 116
 DB 569 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCCTTTTG 624

RESULT 9
 EX464579 853 bp mRNA linear EST 22-MAY-2003
 LOCUS BX464579 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
 DEFINITION clone CS0D003YB14 5-PRIME, mRNA sequence.
 ACCESSION BX464579
 VERSION BX464579.1 GI:31031641
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 1 (bases 1 to 853)
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by life technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 6672.r for
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS1D001ZF10QPL&cluster=6672.r. Contact :
 Feng Liang Email: fliang@life-tech.com URL :
 http://fulllength.invitrogen.com/ invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS1D001ZF10QPL.
 Location/Qualifiers
 1..853
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D003YB14"
 /issue_type="B CELLS (RAMOS CELL LINE)"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo (dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN

Query Match 97.2%; Score 112.8; DB 13; Length 853;
 Best Local Similarity 98.3%; Pred. No. 5.1e-26;
 Matches 114; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAATACCTTTGTAGAAAAACCAAGAAAGACTTTCAAGCCTGGG 60
 DB 513 ATTATTTTACTGCTGGAATACCTTTGTAGAAAAACCAAGAAAGACTTTCAAGCCTGGG 572

QY 61 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCCTTTTG 116
 DB 573 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCCTTTTG 628

RESULT 10
 BF975166 942 bp mRNA linear EST 22-JAN-2001
 LOCUS BF975166

DEFINITION 60224465771 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4335639 5',
 mRNA sequence.
 ACCESSION BF975166
 VERSION BF975166.1 GI:12342381
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE NIH-MGC http://nigc.nci.nih.gov/.
 1 (bases 1 to 942)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rcmail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L10C1207 row: a column: 16
 High quality sequence stop: 707.
 Location/Qualifiers
 1..942
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4335639"
 /issue_type="Primary B-cells from tonsils (cell line)"
 /lab_host="pDH10B (phage-resistant)"
 /clone_lib="NIH_MGC_48"
 /note="Organ: B-cells; Vector: pOTB7; Site: 1; XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald W. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 97.2%; Score 112.8; DB 10; Length 942;
 Best Local Similarity 98.3%; Pred. No. 5.2e-26;
 Matches 114; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATTATTTTACTGCTGGAATACCTTTGTAGAAAAACCAAGAAAGACTTTCAAGCCTGGG 60
 DB 506 ATTATTTTACTGCTGGAATACCTTTGTAGAAAAACCAAGAAAGACTTTCAAGCCTGGG 565

QY 61 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCCTTTTG 116
 DB 566 AAGGCTGCATGAAATTCAGTTCCTCCAGACAGCTTCGGCGCATCCTTTTG 621

RESULT 11
 BX402063 1201 bp mRNA linear EST 13-MAY-2003
 LOCUS BX402063 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 DEFINITION Homo sapiens cDNA clone CS0D012YD18 5-PRIME, mRNA sequence.
 ACCESSION BX402063
 VERSION BX402063.1 GI:30626645
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 1 (bases 1 to 1201)
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)

COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6672.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DL012DB09QPI&cluster=6672.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DL012DB09QPI.

FEATURES

SOURCE

1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL012YD18"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_id="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match

Best Local Similarity 89.1%; Score 103.4; DB 13; Length 1201;
Matches 115; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY

1 ATTATTTTACTGCTGGAATCT-TTTGTAGAAACACGAAAGACTTCAAGCCGTG 59
|||||

Db

579 ATTATTTTACTGCTGGAATCTTTGTAGAAACACGAAAGACTTCAAGCCGTG 638
|||||

QY

60 GAAGGCTGCATGAATAATTCAGTTCTCTCCAGACAGCTTGGCGCATCTTTTG 116
|||||

Db

639 GAAGGCTGCATGAATAATTCAGTTCTCTCCAGACAGCTTGGCGCATCTTTTG 695
|||||

RESULT 12

Bg755005

1034 bp mRNA linear EST 15-MAY-2001
60271151F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4851819 5',

DEFINITION

mRNA sequence.

ACCESSION

BG755005

VERSION

BG755005.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 1034)

AUTHORS

NIH-MGC

TITLE

NIH-MGC

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L16M1695 row: e column: 04
High quality sequence stop: 622.
Location/Qualifiers

FEATURES

SOURCE

1..1034
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 86.2%; Score 100; DB 12; Length 1034;
Best Local Similarity 91.4%; Pred. No. 8e-22; Indels 0; Gaps 0;
Matches 106; Conservative 0; Mismatches 10;

QY

1 ATTATTTTACTGCTGGAATCTTTGTAGAAACACGAAAGACTTCAAGCCGTG 60
|||||

Db

45 ATTATTTTACTGCTGGAATCTTTGTAGAAACACGAAAGACTTCAAGCCGTG 104
|||||

QY

61 AAGGCTGCATGAATAATTCAGTTCTCTCCAGACAGCTTGGCGCATCTTTTG 116
|||||

Db

105 AAGGCTGCATGAATAATTCAGTTCTCTCCAGACAGCTTGGCGCATCTTTTG 160
|||||

RESULT 13

BH302559/c

LOCUS

BH302559 442 bp DNA linear GSS 30-NOV-2001
CH230-100C8.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone

DEFINITION

CH230-100C8, genomic survey sequence.

ACCESSION

BH302559

VERSION

BH302559.1

KEYWORDS

GSS.

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

REFERENCE

1 (bases 1 to 442)

AUTHORS

Zhao, S., Shetty, J., Shatman, S., Teague, G., Geer, K.,
Shvartsbeyn, A., Gebregeorgis, E., Overton, L., Russell, D., Chen, D.,
Riggs, P., de Jong, P. and Fraser, C.M.

TITLE

Rat BAC End Sequences from Library CHORI-230 EcORI segment

JOURNAL

Unpublished (1999)

COMMENT

Other GSSs: CH230-100C8.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@igf.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdjong@mail.choi.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/tdb/bac_end/rat/bac_end_intro.html
Plate: 100 row: C column: 8
Seq primer: T7
Class: BAC ends.
Location/Qualifiers

FEATURES

SOURCE

1..442
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-100C8"
/sex="Female"
/cell_type="Brain"

/clone_1lb="CHORI-230 Segment 1"
/note="Vector: pTARBPAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/SENHed/MCW) BAC library produced by
Pieter de Jong"

ORIGIN

Query Match 84.8%; Score 98.4; DB 28; Length 442;
Best Local Similarity 90.5%; Pred. No. 2,4e-21;
Matches 105; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Db 358 ACATATTTTACTGCTGGAATGATCTTTGTAGAAAACCAAGAACTTTCAAGCCTGGG 60

Qy 61 AAGGGCTGATGAAATTCAGTTCCTCCGAGAGAGCTGGCGGCGATCCTTTTG 116
298 AAGGGCTGATGAAATTCAGTTCCTCCGAGAGAGCTGGCGGCGATCCTTTTG 243

RESULT 14

BG144705

LOCUS 522 bp mRNA linear EST 01-FEB-2001
DEFINITION uc73107.y1 Soares mouse NMGB B-cell Mus musculus cDNA clone
IMAGE:333637.5; Similar to TR:Q9WE0 Q9WE0 ACTIVATION-INDUCED
CYTIDINE DEAMINASE.; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

MUS

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT: Robert Straube, Ph.D.

Email: rstraube@nih.gov

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:1077801

Seq primer: -40RP from Gibco

High quality sequence stop: 487.

Location/Qualifiers

1. 522

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:333637"

/lab_host="DH10B (phage-resistant)"

/clone_1lb="Soares mouse NMGB B-cell"

/note="Organ: germinal B-cell; Vector: pT73D-Pac

(pharmacia) with a modified polylinker; Site 1: Not I;

Site 2: Eco RI; 1st strand cDNA was primed with a Not I -

oligo(dT) primer [5

TGTACCAATCATGAGTGGAGGCGCGCTGTTTTTTTTTTTTTTTTTTTTTTTT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified pT73 vector.

Library is normalized; constructed by Bento Soares and

M.Patima Bonaldo."

ORIGIN

Query Match

Best Local Similarity

Matches 104; Conservative

Qy 1

Db 345

Qy 61

AAGGGCTGATGAAATTCAGTTCCTCCGAGAGAGCTGGCGGCGATCCTTTTG 116

Db

RESULT 15

AK080144

LOCUS

DEFINITION

MUS

Mus musculus adult male aorta and vein cDNA, RIKEN full-length

enriched library, clone:A530070C03 product:activation-induced

cytidine deaminase, full insert sequence.

AK080144

VERSION

KEYWORDS

SOURCE

ORGANISM

MUS

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

405 AAGGGCTGATGAAATTCAGTTCCTCCGAGAGAGCTGGCGGCGATCCTTTTG 460

AK080144 1292 bp mRNA linear HTC 19-SEP-2003
MUS musculus adult male aorta and vein cDNA, RIKEN full-length
enriched library, clone:A530070C03 product:activation-induced
cytidine deaminase, full insert sequence.

AK080144.1 GI:26099024
HTC; CAP trapper.

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

10349636

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

20499374

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, U., Nishi, K., Kikunari, T., Tashiro, H., Itoh, M.,
Sun, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

20530913

The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

11076861

The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

6 (bases 1 to 1292)

Adachi, J., Aizawa, K., Akiyama, T., Aizawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuta, M., Hanagaki, T., Harada, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirose, T., Hirose, T.,
Hori, F., Imocari, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishii, K., Nomura, K., Numata, K., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazawa, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

COMMENT

Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN,
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>.
Location/Qualifiers

FEATURES

SOURCE

1..1292
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57Bl/6J"
/db_xref="FANTOM_DB:A530070C03"
/db_xref="MGI:2405961"
/db_xref="taxon:10090"
/clone="A530070C03"
/sex="male"
/tissue_type="aorta and vein"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1..1292
/note="activation-induced cytidine deaminase
(MGI:MGI:1342279, GB|NM_009645, evidence: BLASTN, 99%,
match=1048)"

ORIGIN

Query Match 83.4%; Score 96.8; DB 11; Length 1292;
Best Local Similarity 89.7%; Pred. No. 9.1e-21;
Matches 104; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1 ATTATTTTACTGCTGGAATCTTTGTAGAAAACCGAAGAACTTCAAGCCTGGG 60
DB 520 ACTATTTTACTGCTGGAATCAATTTGTAGAAAATCGTGAAGAACTTCAAGCCTGGG 579
QY 61 AAGGCTGCATGAATAATTGCTGCTCTCCAGACAGCTTCGGCGCATCCTTTTG 116
DB 580 AAGGCTACATGAATAATTGCTGCTCTCCAGACAGCTTCGGCGCATCCTTTTG 635

Search completed: March 13, 2004, 00:39:04
Job time : 370.498 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 12:27:59 / Search time 6938.04 Seconds

(without alignments)
13767.261 Million cell updates/sec

Title: US-09-966-880A-15

Perfect score: 2172

Sequence: 1 ccccgatagaggttgatga.....agatcagcgataaacttg 2172

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_srs:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_mu:*

19: em_mu:*

20: em_mu:*

21: em_or:*

22: em_ov:*

23: em_pac:*

24: em_pb:*

25: em_pl:*

26: em_ro:*

27: em_srs:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inh:*

32: em_htg_inh:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rtd:*

36: em_htg_rtd:*

37: em_htg_vrt:*

38: em_sy:*

39: em_hugo_hum:*

40: em_hugo_mus:*

41: em_hugo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1 | 2172 | 100.0 | 2172 | 6 | BD016840 |
| 2 | 2172 | 100.0 | 2791 | 6 | AB040431 |
| 3 | 2172 | 100.0 | 2818 | 6 | BD016833 |
| 4 | 2172 | 100.0 | 6564 | 6 | BD016835 |
| 5 | 2172 | 100.0 | 11204 | 6 | BD016860 |
| 6 | 2172 | 100.0 | 11204 | 9 | AB040430 |
| 7 | 2172 | 100.0 | 71132 | 9 | AC092184 |
| 8 | 1199.6 | 55.2 | 1828 | 9 | BC006296 |
| 9 | 307 | 14.1 | 173071 | 9 | AC009477 |
| 10 | 305.6 | 14.1 | 123101 | 9 | HS108F13 |
| 11 | 304.6 | 14.0 | 135404 | 9 | AC084024 |
| 12 | 304.6 | 14.0 | 174492 | 2 | AC024242 |
| 13 | 304.6 | 14.0 | 195623 | 2 | AP004218 |
| 14 | 303.6 | 14.0 | 172884 | 2 | AC091164 |
| 15 | 303.6 | 14.0 | 176063 | 2 | AC024035 |
| 16 | 303.6 | 14.0 | 194091 | 9 | AP003534 |
| 17 | 303.6 | 14.0 | 195323 | 2 | AP003549 |
| 18 | 303 | 14.0 | 159692 | 2 | AC013693 |
| 19 | 303 | 14.0 | 161806 | 2 | AC016495 |
| 20 | 303 | 14.0 | 183859 | 9 | AC023632 |
| 21 | 302.8 | 13.9 | 131943 | 9 | AC005484 |
| 22 | 301.4 | 13.9 | 183586 | 9 | AC092562 |
| 23 | 301 | 13.9 | 234481 | 2 | AC022214 |
| 24 | 301 | 13.9 | 235968 | 9 | AC087393 |
| 25 | 299.4 | 13.8 | 163520 | 9 | AL391259 |
| 26 | 298.8 | 13.8 | 183461 | 9 | AC006121 |
| 27 | 298.2 | 13.7 | 180290 | 9 | AC005081 |
| 28 | 298.2 | 13.7 | 204679 | 9 | AC099676 |
| 29 | 297.6 | 13.7 | 111257 | 9 | AL354917 |
| 30 | 297.6 | 13.7 | 147885 | 2 | AL353793 |
| 31 | 297.4 | 13.7 | 130855 | 9 | AC004089 |
| 32 | 297.4 | 13.7 | 219935 | 9 | AC005015 |
| 33 | 297.2 | 13.7 | 119794 | 9 | HS295844 |
| 34 | 297.2 | 13.7 | 193608 | 2 | AC019126 |
| 35 | 296.2 | 13.6 | 179372 | 2 | AC040922 |
| 36 | 295.6 | 13.6 | 101722 | 9 | AC132192 |
| 37 | 295.2 | 13.6 | 190076 | 9 | AC008403 |
| 38 | 294.6 | 13.6 | 137506 | 9 | AC004824 |
| 39 | 294 | 13.5 | 143773 | 9 | AC005631 |
| 40 | 294 | 13.5 | 153125 | 9 | AC104456 |
| 41 | 293.8 | 13.5 | 112386 | 9 | AL138812 |
| 42 | 293.8 | 13.5 | 132987 | 9 | AC026970 |
| 43 | 293.8 | 13.5 | 153652 | 2 | AC092767 |
| 44 | 293.8 | 13.5 | 159439 | 9 | AC074008 |
| 45 | 293.8 | 13.5 | 169886 | 2 | AC140007 |

ALIGNMENTS

RESULT 1

BD016840

LOCUS BD016840

DEFINITION Novel cytidine deaminase.

ACCESSION BD016840

VERSION BD016840.1 GI:22558016

KEYWORDS UP 2001245669-A/13.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2172)

AUTHORS Honjo, T. and Muramatsu, M.

TITLE Novel cytidine deaminase

JOURNAL Patent: JP 2001245669-A 13 11-SEP-2001;

| | | | | |
|---------------------------|--|---------|--|-----------------|
| Dd | | 1801 | AGAGTTAAATGTACTGACGATTCACTTATAATTTAATTAATTTGGCGTCMATG | 1867 |
| Oy | | 1861 | ATTTTTTNTTAAcAgAtttcttTtTGArntGAAnTGAGrccTCAAgttcAcTA | 1920 |
| Dd | | 1861 | ATTTTTTNTVAAcAGAttTCCtTTTCGATATATnTGAAAAGGrCtCCAAGctTCATA | 1920 |
| Oy | | 1921 | AATTAFACCTTAPAAAAngATTCtTAAPAACAgrTAGTAAATGVACATGCAGTAT | 1980 |
| Dd | | 1921 | AAITTTAACTTTAAPAATgATTCtTAAPAACAgrTAGTAAATGVACATGCAGTAT | 1980 |
| Oy | | 1961 | GGrGCTACGAAGCcATTCTCTTGATTTTTnGTAAAcCTTTnTGACAGcAAATTTgCTC | 2040 |
| Dd | | 1961 | GGTGCTACGAAGCcATTCTCTTGATTTTTnGTAAAcCTTTnTGACAGcAAATTTgCTC | 2040 |
| Oy | | 2041 | TGGCTCACCTTCATnCAGTTAAATPAAATgATDAATATTTTGAAAGCTGTAGAANTAA | 2100 |
| Dd | | 2101 | TGCCCTACCTTCcAACcAGTTAAATPAAATgATDAATATTTTGAAAGCTGTAGAATAA | 2100 |
| Oy | | 2101 | TACCcAAATPAAATVAATPAAAGTGATTTATGTAGAGTTAAATPAAAcCAGTATGAT | 2160 |
| Dd | | 2161 | GGATPAAAcCTTG 2172 | |
| Oy | | 2161 | GGATPAAAcCTTG 2172 | |
| Dd | | 2161 | GGATPAAAcCTTG 2172 | |
| RESULT 2 LOCUS | ABO40431 | 2791 bp | mRNA linear | PRI 03-OCT-2000 |
| DEFINITION | Homo sapiens AID mRNA for activation-induced cytidine deaminase, complete CDS. | | | |
| ACCESSION | ABO40431 | | | |
| VERSION | ABO40431.1 GI:9988409 | | | |
| KEYWORDS | AID; activation-induced cytidine deaminase; Human AID. Homo sapiens (human) | | | |
| SOURCE | Homo sapiens | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (sites) Muto,T., Muramatsu,M., Taniwaki,M., Kinoshita,K. and Honjo.T. Isolation, tissue distribution, and chromosomal localization of the human activation-induced cytidine deaminase (AID) gene Genomics 68 (1), 85-88 (2000) | | | |
| JOURNAL MEDLINE PUBLISHED | 20408890 | | | |
| REFERENCE | 10950930 | | | |
| AUTHORS | 2 (sites) Reyv,P., Muto,T., Levy,Y., Geissmann,F., Pieban,A., Sarai,O., Catalan,N., Forgeville,W., Ducicq-Lagehouse,R., Gennery,A., Tecsan,I., Essoy,F., Kayserili,H., Ugazio,A.G., Bronse,N., Muramatsu.M., Notarangelo,L.D., Kinoshita.K., Honjo.T., Fischer,A. and Durandy.A. | | | |
| TITLE | Activation-induced cytidine deaminase (Aid) deficiency causes the autosomal recessive form of the Hyper-Igm syndrome (HIGM2) Cell 102 (5), 565-575 (2000) | | | |
| JOURNAL MEDLINE PUBLISHED | 20460541 | | | |
| REFERENCE | 11007475 | | | |
| AUTHORS | 3 (bases 1 to 2791) Muto,T., Muramatsu,M., Taniwaki,M., Kinoshita,K. and Honjo.T. Direct Submision Submitted (18-MAR-2000) Tasuku Honjo, Kyoto University, Department of Medical Chemistry, Faculty of Medicine; Yoshida, Sakyo-Ku, Kyoto, Kyoto 606-8501, Japan (E-mail:honjotomfour.med.kyoto-u.ac.jp, Tel.:81-75-753-4371(lex.4371), Fax:81-75-753-4386) | | | |
| JOURNAL TITLE | | | | |
| FEATURES | | | | |
| source | location/Qualifiers 1..2791 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" 1.._2791 /gene="AID" 77.._673 /gene="AID" | | | |
| CDS | | | | |

| Query Match | Best Local Similarity | 100.0% | Score 2172 | DB 9 | Length 2791 |
|--------------|-----------------------|--|------------|------|-------------|
| Matches 2172 | Conservative | 0 | Mismatches | 0 | Indels |
| | | 0 | Gaps | 0 | |
| QY | 1 | CCCCGTGATGAGGTGATGACTTACGAGACGCAATTTGCTACTTTGGACCTTGTATAGCAA | 60 | | |
| DB | 620 | CCCCGTGATGAGGTGATGACTTACGAGACGCAATTTGCTACTTTGGACCTTGTATAGCAA | 679 | | |
| QY | 61 | CTTCCAGGAATGTACACACAGATGAATATCTGCTGGAACACGTGATATAAAACAGT | 120 | | |
| DB | 680 | CTTCCAGGAATGTACACACAGATGAATATCTGCTGGAACACGTGATATAAAACAGT | 739 | | |
| QY | 121 | CCCTCAAGTCTTCTGTTTATTTCTTCAACTCTCACTTTCTTAGAGTTTACAGAAAA | 180 | | |
| DB | 740 | CCCTCAAGTCTTCTGTTTATTTCTTCAACTCTCACTTTCTTAGAGTTTACAGAAAA | 799 | | |
| QY | 181 | ATATTTATATACGACTCTTTAAAAAAGATCTATGCTTGAAAAATAGAGAGAACACAGT | 240 | | |
| DB | 800 | ATATTTATATACGACTCTTTAAAAAAGATCTATGCTTGAAAAATAGAGAGAACACAGT | 859 | | |
| QY | 241 | CTGGCCAGGACGCTGCTGCAATTGCTGCAAGTTTGAATGCAACATGTGCTTACTGGGA | 300 | | |
| DB | 860 | CTGGCCAGGACGCTGCTGCAATTGCTGCAAGTTTGAATGCAACATGTGCTTACTGGGA | 919 | | |
| QY | 301 | ATAACAGAACTGCGAGGACCTGGAGCAATCCATAAGTGCAAGTTTTCTATGACTTTTA | 360 | | |
| DB | 920 | ATAACAGAACTGCGAGGACCTGGAGCAATCCATAAGTGCAAGTTTTCTATGACTTTTA | 979 | | |
| QY | 361 | GGTAGAGTAGAGACGAGGATGATCTCTAAAAAGCATGGTAGAGATCAAAATGTTTTTA | 420 | | |
| DB | 980 | GGTAGAGTAGAGACGAGGATGATCTCTAAAAAGCATGGTAGAGATCAAAATGTTTTTA | 1039 | | |
| QY | 421 | TATCAACATCTTTATTTATTTGATTCATTTGAGTTAACAGTGGTGTATGATATGATTTT | 480 | | |
| DB | 1040 | TATCAACATCTTTATTTATTTGATTCATTTGAGTTAACAGTGGTGTATGATATGATTTT | 1099 | | |
| QY | 481 | TCATATCTTTCCCTTGACGTTTACTTCAAGTAAACAACTCTTCATCAGAGCCATGA | 540 | | |
| DB | 1100 | TCATATCTTTCCCTTGACGTTTACTTCAAGTAAACAACTCTTCATCAGAGCCATGA | 1159 | | |
| QY | 541 | TCATATAGACCTCTTAATGAGAGTATCGGTGATTTGACCCCAACCATCTCTCCAA | 600 | | |
| DB | 1160 | TCATATAGACCTCTTAATGAGAGTATCGGTGATTTGACCCCAACCATCTCTCCAA | 1219 | | |
| QY | 601 | GCATTAAATATCAATCATGCGCTGATGTTTAAATCAGCAGAACATGTTTTATGTTTG | 660 | | |
| DB | 1220 | GCATTAAATATCAATCATGCGCTGATGTTTAAATCAGCAGAACATGTTTTATGTTTG | 1279 | | |
| QY | 661 | TACAAAAGAAATGTTATGCGGTGGGATGAGGATATAGACCATGATGCTCACTTCAA | 720 | | |
| DB | 1280 | TACAAAAGAAATGTTATGCGGTGGGATGAGGATATAGACCATGATGCTCACTTCAA | 1339 | | |
| QY | 721 | GCTACTTTAAATTAAGATCTTTAAATGGGACGAGGACTGTGAACAAGACACCTTAATA | 780 | | |
| DB | 1340 | GCTACTTTAAATTAAGATCTTTAAATGGGACGAGGACTGTGAACAAGACACCTTAATA | 1399 | | |
| QY | 781 | TGGGTGATGCTGAGTAGCAAACTCTTCTGAAAAGCAAACTCTTTTAAGAAAGTCCCT | 840 | | |
| DB | 1400 | TGGGTGATGCTGAGTAGCAAACTCTTCTGAAAAGCAAACTCTTTTAAGAAAGTCCCT | 1459 | | |
| QY | 841 | AATTTAGAAACCCCAAACTTACATATCATATTTAGCAAAACATTTGAGAGAAAGTGG | 900 | | |
| DB | 1460 | AATTTAGAAACCCCAAACTTACATATCATATTTAGCAAAACATTTGAGAGAAAGTGG | 1519 | | |

QY 901 CTTGAATGTTGGGAGAGGAAAAATCTATTGGCTCTGTTGGGTCTCTTCATCTCAGAAATG 960
 DB 1520 CTTGAATGTTGGGAGAGGAAAAATCTATTGGCTCTGTTGGGTCTCTTCATCTCAGAAATG 1579
 QY 961 CCAATCAGGTCAGAGGTTTGTCTACATTTTGTATGTGTGTATGTCTCTCCCAAGGTATAT 1020
 DB 1580 CCAATCAGGTCAGAGGTTTGTCTACATTTTGTATGTGTGTATGTCTCTCCCAAGGTATAT 1639
 QY 1021 TAACTATATAGAGAGTTGTGACAAAGAGAAATGATTAAGCTGGGACCGTGGGACAGGC 1080
 DB 1640 TAACTATATAGAGAGTTGTGACAAAGAGAAATGATTAAGCTGGGACCGTGGGACAGGC 1699
 QY 1081 TCAATGTTTACGCTGCTTGGGAGGTTGAGAGAGGAGATGAGCTTGAACAAGTGTTCAA 1140
 DB 1700 TCAATGTTTACGCTGCTTGGGAGGTTGAGAGAGGAGATGAGCTTGAACAAGTGTTCAA 1759
 QY 1141 GGCAGCCTGGGACACATTAACAAGATCTGTCTTCAAAAAAAGAAAAAAGAAA 1200
 DB 1760 GGCAGCCTGGGACACATTAACAAGATCTGTCTTCAAAAAAAGAAAAAAGAAA 1819
 QY 1201 GAGAGAGGCGCGGCGTGGTGGCTCAAGCGCTGTATATCCAGACCTTTGGAGGCGGAGCC 1260
 DB 1820 GAGAGAGGCGCGGCGTGGTGGCTCAAGCGCTGTATATCCAGACCTTTGGAGGCGGAGCC 1879
 QY 1261 GAGAGGATCACCTGTGTGAGAGAGTTTGAACAGAGCTGGCCAGATGGCAAAACCCCGT 1320
 DB 1880 GAGAGGATCACCTGTGTGAGAGAGTTTGAACAGAGCTGGCCAGATGGCAAAACCCCGT 1939
 QY 1321 CTGTACTCAAAATGCAAAATTTAGCAGAGCGTGTAGCAGGCACTGTATATCCAGCTAC 1380
 DB 1940 CTGTACTCAAAATGCAAAATTTAGCAGAGCGTGTAGCAGGCACTGTATATCCAGCTAC 1999
 QY 1381 TTGGAGGCTGAGGAGAGAGATTCGCTTGAACCCAGAGGTGAGAGTTCAGTAACTGCA 1440
 DB 2000 TTGGAGGCTGAGGAGAGAGATTCGCTTGAACCCAGAGGTGAGAGTTCAGTAACTGCA 2059
 QY 1441 GATCGGCGCTGTGACCTCCAGGCTGGGAGCAAGAGCAAGACTGTCTCAAAAAA 1500
 DB 2060 GATCGGCGCTGTGACCTCCAGGCTGGGAGCAAGAGCAAGACTGTCTCAAAAAA 2119
 QY 1501 AAAAAAGAT 1560
 DB 2120 AAAAAAGAT 2179
 QY 1561 TGCAGAGAAATTTGCTTTATCAACAAATGTAAAGAGCCATTAAGGAGATCCCTATTG 1620
 DB 2180 TGCAGAGAAATTTGCTTTATCAACAAATGTAAAGAGCCATTAAGGAGATCCCTATTG 2239
 QY 1621 TCTCTTTGGTGTCTATTTTGTCCCTTAACAACCTGTCTTTGAACAGTGAAGAAAAATTCGA 1680
 DB 2240 TCTCTTTGGTGTCTATTTTGTCCCTTAACAACCTGTCTTTGAACAGTGAAGAAAAATTCGA 2299
 QY 1681 ATAACCAATATCCCTGTGCGGTATTAACCTAGCAACCTTGCAATGAAGATGAGAGATCC 1740
 DB 2300 ATAACCAATATCCCTGTGCGGTATTAACCTAGCAACCTTGCAATGAAGATGAGAGATCC 2359
 QY 1741 ACAGAGAAATCTGAATGCAACATGTCTATTTTAATCTTAATGTGACATAAGTTGTAA 1800
 DB 2360 ACAGAGAAATCTGAATGCAACATGTCTATTTTAATCTTAATGTGACATAAGTTGTAA 2419
 QY 1801 AGAGTTAAATTTGTTACTTCACTGATTCATTTATTTATTTATTTATTTATTTGCTGATG 1860
 DB 2420 AGAGTTAAATTTGTTACTTCACTGATTCATTTATTTATTTATTTATTTGCTGATG 2479
 QY 1861 ATTTTATTAACATGATTTCTTTCTGATATATGAATGAAGAGTTCAGAGCTTCA 1920
 DB 2480 ATTTTATTAACATGATTTCTTTCTGATATATGAATGAAGAGTTCAGAGCTTCA 2539
 QY 1921 AATTTATTAACATGATTTCTTAATTAACAACGTATGATTAATGTAATGCAATGCAATAT 1980
 DB 2540 AATTTATTAACATGATTTCTTAATTAACAACGTATGATTAATGTAATGCAATGCAATAT 2539

QY 1981 GGTGTACGAGAGCCATTTCTTGATTTTAGTAACTTTATGACAGCAAAATTTGCTTC 2040
 DB 2600 GGTGTACGAGAGCCATTTCTTGATTTTAGTAACTTTATGACAGCAAAATTTGCTTC 2659
 QY 2041 TGGCTCACTTTCATCACTTAATTAATGATTAATTTTGGAGCTGTGAAGATTA 2100
 DB 2660 TGGCTCACTTTCATCACTTAATTAATGATTAATTTTGGAGCTGTGAAGATTA 2719
 QY 2101 TACCAATTAATTAATTAATTAATGATTAATTTATGAGATTAATTAATTAATTAATGAT 2160
 DB 2720 TACCAATTAATTAATTAATTAATGATTAATTTATGAGATTAATTAATTAATTAATGAT 2779
 QY 2161 GGAATTAACCTTG 2172
 DB 2780 GGAATTAACCTTG 2791

RESULT 3
 LOCUS BD016833
 DEFINITION Novel cytidine deaminase.
 ACCESSION BD016833
 VERSION BD016833.1 GI:22558009
 KEYWORDS JP 2001245669-A/6.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2818)
 AUTHORS Honjo, T. and Muramatsu, M.
 TITLE Novel cytidine deaminase
 JOURNAL Patent: JP 2001245669-A 6 11-SEP-2001;
 JAPAN TOBACCO INC, TASUKU HONJO
 OS Homo sapiens (human)
 FN JP 2001245669-A/6
 PD 11-SEP-2001
 PF 28-MAR-2000 JP 2000092981
 PI TASUKU HONJO, MASAMICHI, MURAMATSU
 PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12, A61P17/00,
 PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC C12N1/21,
 PC C12N5/10, C12N9/78, C12P21/02, C12P21/08// (C12N1/21, C12N1/19), PC (C12N5/10, C12R1/91), C12N15/00, C12N5/00, (C12N5/00, C12R1/91) CC
 FH Key
 FT 5'UTR (1), (79)
 FT CDS (80), (676)
 FT 3'UTR (677), (2818).
 FEATURES
 source location/Qualifiers
 1..2818
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 100.0%; Score 2172; DB 6; Length 2818;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCTATAGAGGTGATGATCTTACGAGCGGATTTGCTACTTTGGAGCTTTGATAGCA 60
 DB 623 CCCCTATAGAGGTGATGATCTTACGAGCGGATTTGCTACTTTGGAGCTTTGATAGCA 682
 QY 61 CTCGAGATGTCACACAGATGAATATCTCTGCTGAGACAGTGAATAAAAACAGT 120
 DB 683 CTCGAGATGTCACACAGATGAATATCTCTGCTGAGACAGTGAATAAAAACAGT 742
 QY 121 CTTCAAGCTTCTCTGTTTATTTTATTTCACTCACTTTCTTAAGAGTTACAGAAAA 180
 DB 743 CTTCAAGCTTCTCTGTTTATTTTATTTCACTCACTTTCTTAAGAGTTACAGAAAA 802
 QY 181 AATTTATATAGAGCTTTTAAAAAGATTAATGCTTGAATAAGAGAGAGACAGCT 240

Db 803 ATATTATATAGAGCTCTTTAAAAAGATCTATGCTTGAATAATAGAGAGGACAGAGT 862
 Qy 241 CTGGCAGAGGACGCTGCTCAATTTGGTGAAGTTTGAATGCAACCTTGCCCTGACGGA 300
 Db 863 CTGGCAGAGGACGCTGCTCAATTTGGTGAAGTTTGAATGCAACCTTGCCCTGACGGA 922
 Qy 301 ATAAAGAGAGCTGAGAGGACCTGGAGACATCTTAAAGTCAAGTTTTCTATGACTTTTA 360
 Db 923 ATAAAGAGAGCTGAGAGGACCTGGAGACATCTTAAAGTCAAGTTTTCTATGACTTTTA 982
 Qy 361 GGTAGAGTGAAGAGAGAGTATCTTAAAAAGCATGGTGAAGAGATCAATATGTTTTTA 420
 Db 983 GGTAGAGTGAAGAGAGAGTATCTTAAAAAGCATGGTGAAGAGATCAATATGTTTTTA 1042
 Qy 421 TATCAACATCCTTTATTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTT 480
 Db 1043 TATCAACATCCTTTATTTATTTGATTTGATTTGATTTGATTTGATTTGATTTT 1102
 Qy 481 TCTATCTTTTCCCTTGAAGTTTACTTTCAGTAAACAAACTCTTCCATGAGGATGA 540
 Db 1103 TCTATCTTTTCCCTTGAAGTTTACTTTCAGTAAACAAACTCTTCCATGAGGATGA 1162
 Qy 541 TCTATGAGACCTCTTATGAGAGTATCTGGGTGATTTGATGAGAGGATCTCTCCAA 600
 Db 1163 TCTATGAGACCTCTTATGAGAGTATCTGGGTGATTTGATGAGAGGATCTCTCCAA 1222
 Qy 601 GCATTATATCCCAATCATGCGCTGATGTTTTTATCAGAGAGAGATGTTTTTATGTTG 660
 Db 1223 GCATTATATCCCAATCATGCGCTGATGTTTTTATCAGAGAGAGATGTTTTTATGTTG 1282
 Qy 661 TCAAAAGAGAGTTGTTATGGGTGGGAGTGAAGGATGATGATGATGATGATGATGATG 720
 Db 1283 TCAAAAGAGAGTTGTTATGGGTGGGAGTGAAGGATGATGATGATGATGATGATGATG 1342
 Qy 721 GCTACTTTATTAAGGATCTTAAAAAGGAGAGAGAGATGATGATGATGATGATGATG 780
 Db 1343 GCTACTTTATTAAGGATCTTAAAAAGGAGAGAGATGATGATGATGATGATGATGATG 1402
 Qy 781 TGGGTGATGCTGTAAGTGAAGTCTTGTGAAAGCAACTCTTTTAAAGAGAGTCCCT 840
 Db 1403 TGGGTGATGCTGTAAGTGAAGTCTTGTGAAAGCAACTCTTTTAAAGAGAGTCCCT 1462
 Qy 841 AATTAGAAACACCCCAAACTTCAATATCATATTAATGCAAAAGATGGAAGAGTTG 900
 Db 1463 AATTAGAAACACCCCAAACTTCAATATCATATTAATGCAAAAGATGGAAGAGTTG 1522
 Qy 901 CTTGAATGTTGGGAGAGAGAGATCTATGCTCTGCTGCTCTCTCTCTCTCTCTCTCTCA 960
 Db 1523 CTTGAATGTTGGGAGAGAGAGATCTATGCTCTGCTGCTCTCTCTCTCTCTCTCTCTCA 1582
 Qy 961 CCAATCAGGTCAAGGTTGCTCAATTTGTATGTTGATGCTCTCTCTCTCTCTCTCTCTCT 1020
 Db 1583 CCAATCAGGTCAAGGTTGCTCAATTTGTATGTTGATGCTCTCTCTCTCTCTCTCTCTCT 1642
 Qy 1021 TAACTATTAAGAGAGTGTGACAAAGCAATGATTAAGCTGCGAACCGTGGCACGCG 1080
 Db 1643 TAACTATTAAGAGAGTGTGACAAAGCAATGATTAAGCTGCGAACCGTGGCACGCG 1702
 Qy 1081 TCAATGTTTACGCTGCTGGAGAGTTGAGAGAGAGAGATGCTTGAACACAGTGTTCAA 1140
 Db 1703 TCAATGTTTACGCTGCTGGAGAGTTGAGAGAGAGAGATGCTTGAACACAGTGTTCAA 1762
 Qy 1141 GGCACACCTGGGCAATTAACAAGATCTGCTCTCAAAAAAAGAGAGAGAGAGAGAGAG 1200
 Db 1763 GGCACACCTGGGCAATTAACAAGATCTGCTCTCAAAAAAAGAGAGAGAGAGAGAGAG 1822
 Qy 1201 GAGAGAGAGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 Db 1823 GAGAGAGAGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1882
 Qy 1261 GAGCGGATCACTGTGCTGAGAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 Db 1883 GAGCGGATCACTGTGCTGAGAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1942

Qy 1321 CTGTACTCAAAATGCAAAATTTAGCCAGGCGTGTAGAGAGAGAGAGAGAGAGAGAGAG 1380
 Db 1943 CTGTACTCAAAATGCAAAATTTAGCCAGGCGTGTAGAGAGAGAGAGAGAGAGAGAGAG 2002
 Qy 1381 TTGGAGGCTGAGAGAGAGAGATCGCTTGAACCCAGAGAGTGAAGTTGCAAGTGAAG 1440
 Db 2003 TTGGAGGCTGAGAGAGAGAGATCGCTTGAACCCAGAGAGTGAAGTTGCAAGTGAAG 2062
 Qy 1441 GATCGTCCGCTGCACTCCAGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
 Db 2063 GATCGTCCGCTGCACTCCAGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2122
 Qy 1501 AAAAAAG 1560
 Db 2123 AAAAAAG 2182
 Qy 1561 TGCAGAGAAATGCTTATTCACAAATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
 Db 2183 TGCAGAGAAATGCTTATTCACAAATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2242
 Qy 1621 TCTCTTTGCTGCTATTTGCTCCCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
 Db 2243 TCTCTTTGCTGCTATTTGCTCCCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2302
 Qy 1681 ATAAAGAAATCCCTGAGCGGTTATTACCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
 Db 2303 ATAAAGAAATCCCTGAGCGGTTATTACCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2362
 Qy 1741 ACAGAGAAATCTGAAG 1800
 Db 2363 ACAGAGAAATCTGAAG 2422
 Qy 1801 AGAGTAAAAATGTTACTTATCATGATATCATTTATTTATTTATTTATTTATTTATTT 1860
 Db 2423 AGAGTAAAAATGTTACTTATCATGATATCATTTATTTATTTATTTATTTATTTATTT 2482
 Qy 1861 ATTTTATTTATTAAGATTTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1920
 Db 2483 ATTTTATTTATTAAGATTTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2542
 Qy 1921 AATTATTAATTAAGATTTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980
 Db 2543 AATTATTAATTAAGATTTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2602
 Qy 1981 GGTGCTACAGAGCATTTCTCTGATTTTATGTAACCTTTATGACAGCAAAATTTGCTTC 2040
 Db 2603 GGTGCTACAGAGCATTTCTCTGATTTTATGTAACCTTTATGACAGCAAAATTTGCTTC 2662
 Qy 2041 TGGCTCACTTCAATCAGTAAATTAATGATTAATTTTGAAGAGAGAGAGAGAGAGAGAG 2100
 Db 2663 TGGCTCACTTCAATCAGTAAATTAATGATTAATTTTGAAGAGAGAGAGAGAGAGAGAG 2722
 Qy 2101 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160
 Db 2723 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2782
 Qy 2161 GGAATTAACCTG 2172
 Db 2783 GGAATTAACCTG 2794

RESULT 4
 BD016835
 LOCUS BD016835
 DEFINITION Novel cytidine deaminase.
 ACCESSION BD016835.1 GI:22558011
 VERSION JP 2001245669-A/8.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Db 5481 ACAGGAACTGATGACAACTGCTGATTTTATCTTATGATGACATTAAGTTGTAA 5540
QY 1801 AGAGTAAATGTTACTTCAATGATGATGATTAATTTATTTATTTTGGCTAAAG 1860
Db 5541 AGAGTAAATGTTACTTCAATGATGATGATTAATTTATTTTGGCTAAAG 5600
QY 1861 ATTTTATTAACATGATTTCTTTCTGATATATTGAATGAGCTCAAGCTTCATA 1920
Db 5601 ATTTTATTAACATGATTTCTTTCTGATATATTGAATGAGCTCAAGCTTCATA 5660
QY 1921 AATTATTAACATGATTTCTTTCTGATATATTGAATGAGCTCAAGCTTCATA 1980
Db 5661 AATTATTAACATGATTTCTTTCTGATATATTGAATGAGCTCAAGCTTCATA 5720
QY 1981 GGTGCTACGAGGCAATTTCTTTCTGATATATTGAATGAGCTCAAGCTTCATA 2040
Db 5721 GGTGCTACGAGGCAATTTCTTTCTGATATATTGAATGAGCTCAAGCTTCATA 5780
QY 2041 TGGCTCACTTCAATGATGATTAATGATTAATTTTGAAGCTGTGAAGATTAATA 2100
Db 5781 TGGCTCACTTCAATGATGATTAATGATTAATTTTGAAGCTGTGAAGATTAATA 5840
QY 2101 TACCAATTAATTAATTAATTAATGATTTATGAAATTAATTAATTAATGATGAT 2160
Db 5841 TACCAATTAATTAATTAATTAATGATTTATGAAATTAATTAATTAATGATGAT 5900
QY 2161 GGAATAAAGTTG 2172
Db 5901 GGAATAAAGTTG 5912

RESULT 5
LOCUS BD016860 11204 bp DNA linear PART 27-AUG-2002
DEFINITION Novel cytidine deaminase.
ACCESSION BD016860
VERSION BD016860.1 GI:22558036
KEYWORDS JP 2001245669-A/33.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 11204)
AUTHORS Honjo, T. and Muramatsu, M.
TITLE Novel cytidine deaminase
JOURNAL Patent: JP 2001245669-A 33 11-SEP-2001;
JAPAN TOBACCO INC, TASUKU HONJO
COMMENT OS Homo sapiens (human)
PN JP 2001245669-A/33
PD 11-SEP-2001
PI 28-MAR-2000 JP 2000092981
PI TASUKU HONJO, MASAMICHI MURAMATSU
PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12,
PC A61P17/00
PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12N9/78, C12P21/02, C12P21/08// (C12N1/21, C12R1/19), PC
(C12N5/10, C12R1/91), C12N15/00, C12N5/00, (C12N5/00, C12R1/91) CC
FH Key Location/Qualifiers
1. 11204
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 2172; DB 6; Length 11204;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCCCTATGAGGTGATGATGATTAAGAGCGAATTTCTGACTTTGGGACTTTGATAGCA 60
Db 8956 CCCCCCTATGAGGTGATGATGATTAAGAGCGAATTTCTGACTTTGGGACTTTGATAGCA 9015

QY 61 CTTCCAGAAATGTCACACAGATGAATATCTCTGCTGAAGACAGTGAATAAAAACAGT 120
Db 9016 CTTCCAGAAATGTCACACAGATGAATATCTCTGCTGAAGACAGTGAATAAAAACAGT 9075
QY 121 CTTCAAGCTCTCTGTTTTTATCTTCACTCTCACTTTCTTGAAGATTACAGAAAA 180
Db 9076 CTTCAAGCTCTCTGTTTTTATCTTCACTCTCACTTTCTTGAAGATTACAGAAAA 9135
QY 181 ATATTTATTAACGATCTTTTAAAAAATCTATGCTTTGAAAAATAGAGAAAGAACAGAGT 240
Db 9136 ATATTTATTAACGATCTTTTAAAAAATCTATGCTTTGAAAAATAGAGAAAGAACAGAGT 9195
QY 241 CTGCGCAGGAGCGTGCTGCAATGGTGCAGTTTGAATGCAATGATCCCTACTGGGA 300
Db 9196 CTGCGCAGGAGCGTGCTGCAATGGTGCAGTTTGAATGCAATGATCCCTACTGGGA 9255
QY 301 ATACAGAACTGCAGAACTGGAGCATCTTAAGTGTCAACGTTTTCTATGACTTTTA 360
Db 9256 ATACAGAACTGCAGAACTGGAGCATCTTAAGTGTCAACGTTTTCTATGACTTTTA 9315
QY 361 GGTAGATGAGAGAGAGAGGTGATCTTAATAAAGCATGTGAGAGATCAATGTTTTTA 420
Db 9316 GGTAGATGAGAGAGAGAGGTGATCTTAATAAAGCATGTGAGAGATCAATGTTTTTA 9375
QY 421 TATCAACATCCTTATTAATTTGATGATGATTAAGATTAAGTGTGATGATAGATTTT 480
Db 9376 TATCAACATCCTTATTAATTTGATGATGATTAAGATTAAGTGTGATGATAGATTTT 9435
QY 481 TCTATCTTTCCCTTACAGCTTACTTCAAGTAACAAACCTCTTCAACAGGCCATGA 540
Db 9436 TCTATCTTTCCCTTACAGCTTACTTCAAGTAACAAACCTCTTCAACAGGCCATGA 9495
QY 541 TCTATGAGACTCTTAAAGAGATCTGGGTGATTTGACCCCAACCATCTTCCAAA 600
Db 9496 TCTATGAGACTCTTAAAGAGATCTGGGTGATTTGACCCCAACCATCTTCCAAA 9555
QY 601 GCATTAATATCCATCATGCTGCTGATGTTTATCAGCAGAAAGCATGTTTTATGTTTG 660
Db 9556 GCATTAATATCCATCATGCTGCTGATGTTTATCAGCAGAAAGCATGTTTTATGTTTG 9615
QY 661 TACAAAGAAAGATTTGATGAGTGGAGTAGAGTATAGCAATGATGATGATGATGAT 720
Db 9616 TACAAAGAAAGATTTGATGAGTGGAGTAGAGTATAGCAATGATGATGATGATGAT 9675
QY 721 GCTACTTTAATAAGATCTTAAATGGGAGAGAGACTGTGAACAAGACACCTTATTA 780
Db 9676 GCTACTTTAATAAGATCTTAAATGGGAGAGAGACTGTGAACAAGACACCTTATTA 9735
QY 781 TGGGTGATGCTGGAAGTAGCAATCTTTGAAACGCAAACTCTTTTAAAGAGTCCCT 840
Db 9736 TGGGTGATGCTGGAAGTAGCAATCTTTGAAACGCAAACTCTTTTAAAGAGTCCCT 9795
QY 841 AATTATGAACACCCCAAACTTCAATATCAATTAATGAACAAATTTGAGAGAGTGG 900
Db 9796 AATTATGAACACCCCAAACTTCAATATCAATTAATGAACAAATTTGAGAGAGTGG 9855
QY 901 CTTGAATGTTGGGAGAGAGAAATCTATGAGCTCTGAGGAGTCTTCAATCTCAGAAATG 960
Db 9856 CTTGAATGTTGGGAGAGAGAAATCTATGAGCTCTGAGGAGTCTTCAATCTCAGAAATG 9915
QY 961 CCAATCAGTCAAGGTTTGTCTACATTTGTATGTGTATGCTTCTCCAAAGGTATAT 1020
Db 9916 CCAATCAGTCAAGGTTTGTCTACATTTGTATGTGTATGCTTCTCCAAAGGTATAT 9975
QY 1021 TAACTATTAAGAGATTTGAGCAAAAACAGATGATTAAGCTTCCGAACCTGGCAACAGC 1080
Db 9976 TAACTATTAAGAGATTTGAGCAAAAACAGATGATTAAGCTTCCGAACCTGGCAACAGC 10035
QY 1081 TCAATGTTTCAAGCTGCTGGAGAGTTGAGAGAGGAGATGAGCTTGAACAAGGTGTTCAA 1140
Db 10036 TCAATGTTTCAAGCTGCTGGAGAGTTGAGAGAGGAGATGAGCTTGAACAAGGTGTTCAA 10095

| | | | | |
|----|--|-------|--|-------|
| OY | | 1141 | GGCCAGCCTGSGCAACAATAACAAAGTCCTGTCTCTCAAAAAAAAAAAAAAAAAAAGAAA | 1200 |
| Dd | | 10096 | GGCACAGCCTGGGCAACATPACAAAGATCTGTCTCTTCAAAAAAAAAAAAAAAAAAGAAA | 10155 |
| OY | | 1201 | GAGAGAGGGCGCGGGCGTGATGGCTCACGCCCTGTATCCAGCATTTGGGAAGCCGAGCC | 1260 |
| Dd | | 10156 | GAGAGAGGGCGCGGGCGTGATGGCTCACGCCCTGTATCCAGCATTTGGGAAGCCGAGCC | 10215 |
| OY | | 1261 | GGGCGGATCACCTGTGGTCAAGAATTGAGACCAAGCTGGCCAACTGGCAAACCCTGT | 1320 |
| Dd | | 10216 | GGGCGGATCACCTGTGGTCAAGAATTGAGACCAAGCTGGCCAACTGGCAAACCCTGT | 10275 |
| OY | | 1321 | CTGTACTCAAAATGCAAAAAATTAGCCAGGGGTGGTAGCAGCAACCTGTATATCCCAAGTAC | 1380 |
| Dd | | 10276 | CTGTACTCAAAATGCAAAAAATTAGCCAGGGGTGGTAGCAGCAACCTGTATATCCCAAGTAC | 10335 |
| OY | | 1381 | TTGGGAGGCTGAGGCGAGAGATCGCTTTGAACCCAGAGGGTGGAGGTTGCATTAAGCTGA | 1440 |
| Dd | | 10336 | TTGGGAGGCTGAGGCGAGAGATCGCTTTGAACCCAGAGGGTGGAGGTTGCATTAAGCTGA | 10395 |
| OY | | 1441 | GATCGTGCCGTTGCACTCCAGCCTGGGCGAACAGACAAGCTGTGCTCAGAAAAAAA | 1500 |
| Dd | | 10396 | GATCGTGCCGTTGCACTCCAGCCTGGGCGAACAGACAAGCTGTGCTCAGAAAAAAA | 10455 |
| OY | | 1501 | AAAAAAAAAG | 1560 |
| Dd | | 10456 | AAAAAAAAAG | 10515 |
| OY | | 1561 | TGCAAGGAAATGATGCTTTATCCAAAGAAATGTAGAGAGCAATTAAGGATCCCTATTG | 1620 |
| Dd | | 10516 | TGCAAGGAAATGATGCTTTATCCAAAGAAATGTAGAGAGCAATTAAGGATCCCTATTG | 10575 |
| OY | | 1621 | TCTCTTTGGTGTATTTTGTCCCTAACCACTGTCTTGAAGTGAAGAAAAATATTCAGA | 1680 |
| Dd | | 10576 | TCTCTTTGGTGTATTTTGTCCCTAACCACTGTCTTGAAGTGAAGAAAAATATTCAGA | 10635 |
| OY | | 1681 | ATAACCATATCCCTGCGCGTTATTAACCTGACCAACCTTGCAATGAAGTAGAGCATCC | 1740 |
| Dd | | 10636 | ATAACCATATCCCTGCGCGTTATTAACCTGACCAACCTTGCAATGAAGTAGAGCATCC | 10695 |
| OY | | 1741 | AACAGAAAACCTTGAAATGACAAACGCTCTTATTTAACTATGTGCATTAAGTTGTAA | 1800 |
| Dd | | 10696 | AACAGAAAACCTTGAAATGACAAACGCTCTTATTTAACTATGTGCATTAAGTTGTAA | 10755 |
| OY | | 1801 | AGAGTAAAAATGTATCTCATGTATTCATTTATTTATTTATTTATTTATTTATTTATTT | 1860 |
| Dd | | 10756 | AGAGTAAAAATGTATCTCATGTATTCATTTATTTATTTATTTATTTATTTATTTATTT | 10815 |
| OY | | 1861 | ATTTTTATTAACATGATTTCTTTTCGTATATATTGAATGGAAGTCTCAAGCTTCATA | 1920 |
| Dd | | 10816 | ATTTTTATTAACATGATTTCTTTTCGTATATATTGAATGGAAGTCTCAAGCTTCATA | 10875 |
| OY | | 1921 | AATTTATACCTTTAGAAATGATTTCTAATTAACAAGATATGAATTTGTAACATTTGCATAT | 1980 |
| Dd | | 10876 | AATTTATACCTTTAGAAATGATTTCTAATTAACAAGATATGAATTTGTAACATTTGCATAT | 10935 |
| OY | | 1981 | GGTGCTAAGAGCCATTTCTGTATTTTATAGTAACTTTTATGACAGCAAAATTTGCTTC | 2040 |
| Dd | | 10936 | GGTGCTAAGAGCCATTTCTGTATTTTATAGTAACTTTTATGACAGCAAAATTTGCTTC | 10995 |
| OY | | 2041 | TGGCTCATCTTCAATCAGTTAAATTAATGAATTAATTTTGAAGCTGTGAAGATPAAA | 2100 |
| Dd | | 10996 | TGGCTCATCTTCAATCAGTTAAATTAATGAATTAATTTTGAAGCTGTGAAGATPAAA | 11055 |
| OY | | 2101 | TACCAATPAATTAATTAATTAAGATTTATGAAGTAAATTAATTAATTAATTAATTAATTA | 2160 |
| Dd | | 11056 | TACCAATPAATTAATTAATTAAGATTTATGAAGTAAATTAATTAATTAATTAATTAATTA | 11115 |
| OY | | 2161 | GGAAATPAACCTG 2172 | |
| Dd | | 11116 | GGAAATPAACCTG 11127 | |

[illegible]

repeat_region 5305. .5563
/rpt_family="AluJo"
repeat_region 5564. .5599
/rpt_family="Alu_rich"
repeat_region 5644. .5733
/rpt_family="MSTd"
repeat_region 5734. .6012
/rpt_family="AluSg"
repeat_region 6020. .6336
/rpt_family="AluSx"
repeat_region 6337. .6629
/rpt_family="MSTd"
repeat_region 6630. .6650
/rpt_family="CAA)n"
repeat_region 6711. .6817
/rpt_family="MIR"
repeat_region 7033. .7327
/rpt_family="AluSx"
repeat_region 7607. .7745
/rpt_family="FLAM_A"
repeat_region 7980. .8047
/rpt_family="MIR"
repeat_region 8050. .8549
/rpt_family="LTR47A"
repeat_region 8567. .8698
/rpt_family="MIR"
repeat_region 8815. .9101
/rpt_family="AluSg"
repeat_region 11227. .12175
/rpt_family="LTR5"
STS 12625. .12900
/standard_name="87432"
repeat_region 13547. .13664
/rpt_family="LIMB5"
repeat_region 13665. .13979
/rpt_family="AluY"
repeat_region 13980. .14193
/rpt_family="LIMB5"
repeat_region 14622. .14924
/rpt_family="AluY"
repeat_region 15813. .17136
/rpt_family="LIPAB"
repeat_region 17272. .17395
/rpt_family="FLAM_C"
repeat_region 17426. .17575
/rpt_family="MIR"

Query Match 100.0%; Score 2172; DB 9; Length 71132;

Best Local Similarity 100.0%; Pred. No. 0; Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCCCTGATGAGGTGATGACTTACGAGCGCATTCCTGACTTTGGACTTTGATGACAA 60
43499 CCCCTGATGAGGTGATGACTTACGAGCGCATTCCTGACTTTGGACTTTGATGACAA 43558
61 CTTCCAGGAATGTCAACACGATGAATATCTCTGCTGAGACAGTGTGATAAAAACAGT 120
43559 CTTCCAGGAATGTCAACACGATGAATATCTCTGCTGAGACAGTGTGATAAAAACAGT 43618
121 CCTTAAGTCTCTCTGTTTATTTCTTCACTCTCACTTTCTTGAAGTTACGAAAAA 180
43619 CCTTAAGTCTCTCTGTTTATTTCTTCACTCTCACTTTCTTGAAGTTACGAAAAA 43678
181 ATATTATATACGACTCTTTTAAAAAGATCTATGCTTGAATATGAGAGGACACAGGT 240
43679 ATATTATATACGACTCTTTTAAAAAGATCTATGCTTGAATATGAGAGGACACAGGT 43738
241 CTGGCAGGAGAGCTGCTGCAATTTGCTGAGTTTGAATGACAACTTGTCCCTACTGAG 300
43739 CTGGCAGGAGAGCTGCTGCAATTTGCTGAGTTTGAATGACAACTTGTCCCTACTGAG 43798
301 ATAAAGAACTGACGAGCCTGGAGACCTTAAGTGTCAACGTTTCTATGACTTTTA 360

43799 ATAAAGAACTGACGAGCCTGGAGACCTTAAGTGTCAACGTTTCTATGACTTTTA 43858
361 GGTAGATGAGACAGAGGTAGATCCATAAAGACATGGTGAGAGGATCAAAATGTTTTTA 420
43859 GGTAGATGAGACAGAGGTAGATCCATAAAGACATGGTGAGAGGATCAAAATGTTTTTA 43918
421 TATCAACATCCTTATATATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTT 480
43919 TATCAACATCCTTATATATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTT 43978
481 TCTATCTTTTCCCTTGAAGCTTTACTTTCAAGTAAACAACCTTCCATCAGGCCATGA 540
43979 TCTATCTTTTCCCTTGAAGCTTTACTTTCAAGTAAACAACCTTCCATCAGGCCATGA 44038
541 TCTATGAGACCTCTTATAGAGTATCTGAGGATTTGATACCCCAACCATCTCCCAAA 600
44039 TCTATGAGACCTCTTATAGAGTATCTGAGGATTTGATACCCCAACCATCTCCCAAA 44098
601 GCATTATATCAATCATGCGCTGTATGTTTATCATGAGAGACATGTTTATGTTT 660
44099 GCATTATATCAATCATGCGCTGTATGTTTATCATGAGAGACATGTTTATGTTT 44158
661 TACAAAAGAGATTGTTATGAGTGGAGATGAGATGATGATGATGATGATGATGATGAT 720
44159 TACAAAAGAGATTGTTATGAGTGGAGATGAGATGATGATGATGATGATGATGATGAT 44218
721 GCTACTTTTATTAAGGATCTTAAATGGGAGAGAGACTGTGATCAAGACACCTTAATA 780
44219 GCTACTTTTATTAAGGATCTTAAATGGGAGAGAGACTGTGATCAAGACACCTTAATA 44278
781 TGGGTGATGCTGAGAGTGCAGAAATCTGGAAGACCACTTTTAAAGAGAGCCCT 840
44279 TGGGTGATGCTGAGAGTGCAGAAATCTGGAAGACCACTTTTAAAGAGAGCCCT 44338
841 AATTGAAACACCCCAAACTTCAATATCAATATTTAGCAAACTTGAAGAGAGTTG 900
44339 AATTGAAACACCCCAAACTTCAATATTTAGCAAACTTGAAGAGAGAGTTG 44398
901 CTTGAATGTTGGGAGAGAGAGAAATCTATGAGTCTGCTGAGTCTCTCAATCTCAAGAA 960
44399 CTTGAATGTTGGGAGAGAGAGAAATCTATGAGTCTGCTGAGTCTCTCAATCTCAAGAA 44458
961 CCATATAGCTCAAGTTTCTGATCTTTGATGATGATGATGATGATGATGATGATGAT 1020
44459 CCATATAGCTCAAGTTTCTGATCTTTGATGATGATGATGATGATGATGATGATGAT 44518
1021 TAACATATATAGAGATTTGATGACAAACAGATGATTAAGTGCAGAACCGTGCACAGCC 1080
44519 TAACATATATAGAGATTTGATGACAAACAGATGATTAAGTGCAGAACCGTGCACAGCC 44578
1081 TCATAGTCTAGCTGCTTGGAGGTTGAGAGGAGAGATGAGCTTGAACAAGGTTTCAA 1140
44579 TCATAGTCTAGCTGCTTGGAGGTTGAGAGGAGAGATGAGCTTGAACAAGGTTTCAA 44638
1141 GGCACGCTGGGACATATAACAGATCTGCTCTCAAAAAAAAAAAAAAAAAAAAAA 1200
44639 GGCACGCTGGGACATATAACAGATCTGCTCTCAAAAAAAAAAAAAAAAAAAAAA 44698
1201 GAGAGAGGCGCGGCGTGTGCTCAAGCTGTATATCCAGACATTTGGAGAGCGGAGGC 1260
44699 GAGAGAGGCGCGGCGTGTGCTCAAGCTGTATATCCAGACATTTGGAGAGCGGAGGC 44758
1261 GGGCGATACCTGTGATGAGAGTTTGAACAGCTGAGCAATGAGCAAAACCCGCT 1320
44759 GGGCGATACCTGTGATGAGAGTTTGAACAGCTGAGCAATGAGCAAAACCCGCT 44818
1321 CTGTACTCAAAATGCAAAATTTAGCAGCGGTGATGAGAGCACTGTATATCCAGCTAC 1380
44819 CTGTACTCAAAATGCAAAATTTAGCAGCGGTGATGAGAGCACTGTATATCCAGCTAC 44878
1381 TTGGAGGCTGAGGAGAGAAATGCTTGAACCCAGAGGATGAGAGTTGCAATAGCTGTA 1440
44879 TTGGAGGCTGAGGAGAGAAATGCTTGAACCCAGAGGATGAGAGTTGCAATAGCTGTA 44938

QY 1441 GATGTCGCGTTGCACTCCAGCCTGGGCGACAGAGCAAGACTCTGTCTCAGAAAAA 1500
 DB 44939 GATGTCGCGTTGCACTCCAGCCTGGGCGACAGAGCAAGACTCTGTCTCAGAAAAA 44998
 QY 1501 AAAAAAG 1560
 DB 44999 AAAAAAG 45058
 QY 1561 TGCAAGAAATTTGCTTATTCACAAATATGTAAGAGCCATTAAGGAGATCCCTATTG 1620
 DB 45059 TGCAAGAAATTTGCTTATTCACAAATATGTAAGAGCCATTAAGGAGATCCCTATTG 45118
 QY 1621 TCTCTTTGCTGCTATTTGCTTATTCACAACTGCTTTGACAGAGAGAGAGAGAGAG 1680
 DB 45119 TCTCTTTGCTGCTATTTGCTTATTCACAACTGCTTTGACAGAGAGAGAGAGAGAG 45178
 QY 1681 ATTAACCATATTCCTGCTGCTTATTCACAACTGCTTTGACAGAGAGAGAGAGAGAG 1740
 DB 45179 ATTAACCATATTCCTGCTGCTTATTCACAACTGCTTTGACAGAGAGAGAGAGAGAG 45238
 QY 1741 ACAAG 1800
 DB 45239 ACAAG 45298
 QY 1801 AGAGTAAATATTTGCTTATTCACAACTGCTTTGACAGAGAGAGAGAGAGAGAGAG 1860
 DB 45299 AGAGTAAATATTTGCTTATTCACAACTGCTTTGACAGAGAGAGAGAGAGAGAGAG 45358
 QY 1861 ATTTTATTAACATGATTTCTCTGATTTTCTGATATATGTAAGAGAGAGAGAGAGAG 1920
 DB 45359 ATTTTATTAACATGATTTCTCTGATTTTCTGATATATGTAAGAGAGAGAGAGAGAG 45418
 QY 1921 AATTATTAACATTTTGAATATGATTTCTAATAACAGTATGTAACATGTAAGAT 1980
 DB 45419 AATTATTAACATTTTGAATATGATTTCTAATAACAGTATGTAACATGTAAGAT 45478
 QY 1981 GGTGTCAG 2040
 DB 45479 GGTGTCAG 45538
 QY 2041 TGGCTCACTTCAATCACTTCAATCACTTCAATCACTTCAATCACTTCAATCACTTCA 2100
 DB 45539 TGGCTCACTTCAATCACTTCAATCACTTCAATCACTTCAATCACTTCAATCACTTCA 45598
 QY 2101 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160
 DB 45599 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 45658
 QY 2161 GGAATTAATTTG 2172
 DB 45659 GGAATTAATTTG 45670

RESULT 8
 LOCUS BC006296 1828 bp mRNA linear PRI 03-OCT-2003
 DEFINITION Homo sapiens activation-induced cytidine deaminase, mRNA (cdna)
 ACCESSION BC006296
 VERSION BC006296.2 GI:33871601
 KEYWORDS MGC.
 ORGANISM Homo sapiens (human)
 SOURCE
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 1 (bases 1 to 1828)
 Strausberg, R.D., Collins, F.S., Wagner, L., Shenman, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Wang, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, N.J., Uedin, T.B., Toshiyuki, S.,
 Carmichael, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullaly, S.J., Bosak, S.A., Mowbray, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulys, S.M.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahney, U., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzyzanski, M.I., Skalka, U., Smalls, D.E.,
 Schermer, A., Schein, J.E., Jones, S.J. and Marra, M.A.,
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 MEDLINE
 PUBMED
 12477932
 2 (bases 1 to 1828)
 Strausberg, R.
 Direct Submission
 Submitted (09-APR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 On Aug 19, 2003 this sequence version replaced gi:13623400.
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov

Tissue Procurement: Louis Staudt
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILML)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akhtar, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
 Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stancijep, S., Thomas, P.J., Touchman, J.W.,
 Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILML at: <http://imgc.llnl.gov>
 Series: IRAL Plates: 17 Row: a Column: 1
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 10190699.

FEATURES

source

1..1828
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:12911 IMAGE:4054915"
 /tissue_type="Primary B-Cells from Tonsils"
 /clone_id="NIR MGC_48"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1..1828
 /gene="AICDA"
 /note="synonyms: AID, HIGM2, CDA2, ARP2"
 /db_xref="LocusID:57379"
 /db_xref="MIM:605257"
 77..673
 /codon_start=1
 /product="activation-induced cytidine deaminase"
 /protein_id="AAH06296.1"
 /db_xref="GI:13623401"
 /db_xref="LocusID:57379"
 /translation="MDSLLNMRKFLYQKVRWAKGRRETYLCVVRGRDSTFSL
 DGYLRNKGCHVELFLRYSIDMDLDRGCRVYTWFSNPFCDCAAHVADFLRGNP

ORIGIN
NLSRIPTARLYPCEDRKAPEEGLRLRLEHAGVOIAIMTFKDYFCMTFVENHERTEK
AMEGHENSVRSLRQLRILPLVEVDLDRFAETLGL"

Query Match 55.2%; Score 1199.6; DB 9; Length 1828;
Best Local Similarity 99.7%; Pred. No. 3.5e-221;
Matches 1202; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCCGCTATAGAGTTGATGACTTACGAGACGATTCGACTCTTGGAGCTTGGATAGCA 60
DB 620 CCCCTGTATGAGAGTTGATGACTTACGAGACGATTCGACTCTTGGAGCTTGGATAGCA 679
QY 61 CTTCAGAGATGTCAACACGATGAAATATCTGCTGTAAGAAGTGTGATTAATAAAGT 120
DB 680 CTTCAGAGATGTCAACACGATGAAATATCTGCTGTAAGAAGTGTGATTAATAAAGT 739
QY 121 CCTTCAGCTTCTCTGTTTATTTCTTCACTCTGACCTTCTTGAAGTTTACAGAAAA 180
DB 740 CCTTCAGCTTCTCTGTTTATTTCTTCACTCTGACCTTCTTGAAGTTTACAGAAAA 799
QY 181 ATATTATATAGAGCTCTTAAAGATATGCTGTAAGTGTGAAATAGAGAGACAGAGT 240
DB 800 ATATTATATAGAGCTCTTAAAGATATGCTGTAAGTGTGAAATAGAGAGACAGAGT 859
QY 241 CTGGCCAGAGAGCTGCTGCAATTTGGTCAAGTTTGAATGCAACATTTGCCCTACAGG 300
DB 860 CTGGCCAGAGAGCTGCTGCAATTTGGTCAAGTTTGAATGCAACATTTGCCCTACAGG 919
QY 301 ATTAAGAACTGACAGAGCTGGAGAGATCTTAAGTGTCAAGCTTTTCTATGACTTTTA 360
DB 920 ATTAAGAACTGACAGAGCTGGAGAGATCTTAAGTGTCAAGCTTTTCTATGACTTTTA 979
QY 361 GGTAGATAGAGAGAGATGATCTTAAAGATGTTGAGAGATCAATGTTTGA 420
DB 980 GGTAGATAGAGAGAGATGATCTTAAAGATGTTGAGAGATCAATGTTTGA 1039
QY 421 TATCAACATCTTATTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 480
DB 1040 TATCAACATCTTATTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1099
QY 481 TCTATCTTTTCCCTGAGCTTACTTCAAGTAAACAACTCTTCAATCAGGAGATGA 540
DB 1100 TCTATCTTTTCCCTGAGCTTACTTCAAGTAAACAACTCTTCAATCAGGAGATGA 1159
QY 541 TCTATAGAGCTCTTATAGAGATCTGGGATGATTTGATGATCCCAACATCTCTCCAA 600
DB 1160 TCTATAGAGCTCTTATAGAGATCTGGGATGATTTGATGATCCCAACATCTCTCCAA 1219
QY 601 GCAATTAATCCATATAGAGCTGATGATTTTATCAAGAGAGATGTTTATGTTTG 660
DB 1220 GCAATTAATCCATATAGAGCTGATGATTTTATCAAGAGAGATGTTTATGTTTG 1279
QY 661 TCAAAAGAGATGTTTATGAGAGAGATGAGATGATGAGATGATGAGATGATGAGAT 720
DB 1280 TCAAAAGAGATGTTTATGAGAGAGATGAGATGATGAGATGATGAGATGATGAGAT 1339
QY 721 GCTACTTTTATTAAGAGATCTTAAATGGGACAGAGAGATGATGAGATGATGAGAT 780
DB 1340 GCTACTTTTATTAAGAGATCTTAAATGGGACAGAGAGATGATGAGATGATGAGAT 1399
QY 781 TGGGTTGATGCTGAAGTGAATCTTTGAGAAAGCAACTCTTTTAAAGAGTCCCT 840
DB 1400 TGGGTTGATGCTGAAGTGAATCTTTGAGAAAGCAACTCTTTTAAAGAGTCCCT 1459
QY 841 AATTTAGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 900
DB 1460 AATTTAGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1519
QY 901 CTTCGATGTTGGGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 960
DB 1520 CTTCGATGTTGGGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 1579
QY 961 CCAATCAGGTCAAGGTTTGCTAATTTTGTATGTTGTATGTTGTATGTTGTATGTT 1020

DB 1580 CCAATCAGGTCAAGGTTTGCTATCTTATGTTGTATGTTGTATGTTGTATGTTGTAT 1639
QY 1021 TAACTATATAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
DB 1640 TAACTATATAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1699
QY 1081 TCAATGTTTACGCTGTTGGAGAGTTGAGAGAGAGATGATGATGATGATGATGATG 1140
DB 1700 TCAATGTTTACGCTGTTGGAGAGTTGAGAGAGAGATGATGATGATGATGATGATG 1759
QY 1141 GGCAGAGCTGGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
DB 1760 GGCAGAGCTGGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1819
QY 1201 GAGAGA 1206
DB 1820 AAAAAA 1825

RESULT 9
AC009477/c 173071 bp DNA linear PRI 08-NOV-2000
LOCUS Homo sapiens BAC clone RP11-209H16 from 2, complete sequence.
DEFINITION AC009477
ACCESSION AC009477
VERSION AC009477.4 GI:9755488
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 173071)
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PubMed 9847074

REFERENCE 2 (bases 1 to 173071)
Ali,V., Maupin,R. and Doucette,J.
The sequence of Homo sapiens BAC clone RP11-209H16
Unpublished
REFERENCE 3 (bases 1 to 173071)
Waterston,R.H.
Direct Submission
Submitted (24-AUG-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE 4 (bases 1 to 173071)
Waterston,R.H.
Direct Submission
Submitted (09-AUG-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE 5 (bases 1 to 173071)
Waterston,R.
Direct Submission
Submitted (08-NOV-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Aug 9, 2000 this sequence version replaced gi:7630891.
COMMENT
Genome Center
Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.edu
----- Summary Statistics

Center project name: H_NH0209H16

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPc1-11 human BAC library was made from the blood of one male
donor, as described by Osogawa, K., Moon, P.Y., Zhao, B., Frengen, B.,
Tateno, M., Catanesse, J.J. and de Jong, P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genome 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
and coworkers at the Roswell Park Cancer Institute
(<http://bacpac.med.buffalo.edu>)

NEIGHBORING SEQUENCE INFORMATION:

Actual start of this clone is at base position 1 of RPc1-209H16;
actual end is at base position 173071 of RPc1-209H16.

FEATURES

```

source
    1. 173071
       /organism="Homo sapiens"
       /mol_type="genomic DNA"
       /db_xref="taxon:9606"
       /chromosome="2"
       /map="2"
       /clone="RPc1-209H16"
       /clone_1fb="RPc1-11"
       location/Qualifiers
repeat_region
    17..129
       /rpt_family="L1"
repeat_region
    194..338
       /rpt_family="L1"
repeat_region
    343..662
       /rpt_family="Alu"
repeat_region
    707..1014
       /rpt_family="L1"
repeat_region
    1133..1499
       /rpt_family="MALR"
repeat_region
    1761..1825
       /rpt_family="L2"
repeat_region
    1965..2144
       /rpt_family="MALR"
repeat_region
    2148..2578
       /rpt_family="L1"
repeat_region
    2615..2879
       /rpt_family="L1"
repeat_region
    2937..3159
       /rpt_family="MER1_type"
repeat_region
    5170..5465
       /rpt_family="Alu"
repeat_region
    5637..5943
       /rpt_family="Alu"
repeat_region
    8914..8969
       /note="similar to EST W91591 (NID:g1408017)"
repeat_region
    8933..9240
       /rpt_family="Alu"
repeat_region
    1059T..10700
       /rpt_family="MR"
repeat_region
    10736..11707
       /note="CpG island (%GC=74.8, o/e=0.80, #CpGs=113)"
misc_feature
    11160..11360
       /note="similar to EST AW295558 (NID:g6702194)"
misc_feature
    11355..11379
       /note="similar to EST AW449246 (NID:g6990022)"
misc_feature
    11379..11783

```

```

repeat_region
    1882..14181
       /rpt_family="ERV1"
misc_feature
    1494T..15123
       /note="similar to EST AW449246 (NID:g6990022)"
misc_feature
    15463..15761
       /note="similar to EST AW449246 (NID:g6990022)"
misc_feature
    16257..16283
       /note="similar to EST AW956922 (NID:g8146605)"
misc_feature
    16259..16283
       /note="similar to EST AW956922 (NID:g8146605)"
misc_feature
    16352..16531
       /note="similar to EST AW956922 (NID:g8146605)"
misc_feature
    16352..16531
       /note="similar to EST T08962 (NID:g389990)"
misc_feature
    16425..16531
       /note="similar to EST T08962 (NID:g389990)"
misc_feature
    16426..16531
       /note="similar to EST F07143 (NID:g672789)"
misc_feature
    17321..17553
       /note="similar to EST H24095 (NID:g892790) ym50a07.r1"
misc_feature
    17321..17553
       /note="similar to EST AW956922 (NID:g8146605)"
misc_feature
    17321..17553
       /note="similar to EST H24095 (NID:g892790) ym50a07.r1"
misc_feature
    17321..17553
       /note="similar to EST F07143 (NID:g672789)"
misc_feature
    17321..17553
       /note="similar to EST T08962 (NID:g389990)"
misc_feature
    17344..17553
       /note="similar to EST AW12638 (NID:g6098166)"
misc_feature
    17388..17553
       /note="similar to EST AW62170 (NID:g272417)"
misc_feature
    17449..17553
       /note="similar to EST AW62170 (NID:g272417)"
misc_feature
    17475..17553
       /note="similar to EST AW62170 (NID:g272417)"
misc_feature
    17906..18040
       /note="similar to EST AW953277 (NID:g8142960)"
misc_feature
    17906..18036
       /note="similar to EST H24095 (NID:g892790) ym50a07.r1"
misc_feature
    17906..18035
       /note="similar to EST AW415510 (NID:g6943392)"
misc_feature
    17906..18035
       /note="similar to EST AW953277 (NID:g8142960)"
misc_feature
    17906..18035
       /note="similar to EST AW953277 (NID:g8142960)"
misc_feature
    17906..18035
       /note="similar to EST AW956922 (NID:g8146605)"
misc_feature
    17906..18035
       /note="similar to EST AW62170 (NID:g272417)"
misc_feature
    17906..17954
       /note="similar to EST AW12638 (NID:g6098166)"
misc_feature
    17918..18035
       /note="similar to EST H13412 (NID:g878232) y172h12.r1"
repeat_region
    18663..18758
       /rpt_family="MER1_type"
repeat_region
    19059..19343
       /rpt_family="MER1_type"
repeat_region
    19543..19674
       /note="similar to EST AW953277 (NID:g8142960)"
misc_feature
    19543..19674
       /note="similar to EST H13412 (NID:g878232) y172h12.r1"
misc_feature
    19543..19608
       /note="similar to EST M62170 (NID:g272417)"
misc_feature
    19543..19602
       /note="similar to EST AW956922 (NID:g8146605)"
misc_feature
    20388..20633
       /note="similar to EST T80103 (NID:g698612) yd04h06.r1"
misc_feature
    20390..20633
       /note="similar to EST AW953277 (NID:g8142960)"
misc_feature
    20390..20633
       /note="similar to EST H13412 (NID:g878232) y172h12.r1"
misc_feature
    20420..20633
       /note="similar to EST H1836 (NID:g876656) ym15a06.r1"
misc_feature
    21460..21537
       /note="similar to EST H1836 (NID:g876656) ym15a06.r1"
misc_feature
    21460..21517
       /note="similar to EST T80103 (NID:g698612) yd04h06.r1"

```

```

misc_feature      21460..21500
                    /note="similar to EST AW953277 (NTD:G8142960)"
misc_feature      21492..21537

Query Match
Best Local Similarity 78.8%; Pred. No. 2,4e-49;
Matches 379; Conservative 0; Mismatches 100; Indels 2; Gaps 1;

14.1%; Score 307; DB 9; Length 173071;

1048 CAGATGATTAAGCTGCGAAGCGTGGCAGACCGTCAATAGTTCTAGCTGTTGGAGGTTG 1107
136573 CAAAAGACTAAGTTGATGTGTGTGACATACCTGTATCTCAGTACTCGAGGCTG 136514
1108 AGGAGGAGGATGCTTGAACACAGGTTTCAAGGCCAGCTGGGCAATACAAATC 1167
136513 AGCAGAGAGATGCTTGAAGCCAGAGTTTGAACAGCAGCTGGGCAATATATAGACC 136454
1168 CTGTCTCTCAAAAAAAGAAAAAAGAAAAAAGAAAGAGAGGCGCGGCTGTGCTGAC 1227
136453 CTATCTCTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAGCAAGCGCAGTGACTGAG 136394
1228 GCCTGATATCCAGCACTTTGGGAGGCGAGCGCGGCGGATCATCCTGTGTGAGAGTTT 1287
136393 GCGTGAATCCAGCACTTTGGGAGGCGCAAGCGCGGCGGATCATCCTGTGTGAGAGTTT 136334
1288 GAGACCAAGCTGGCACAATGACAAACCCGCTGTGTACTCAAAAATGCAAAATTAGCCA 1347
136333 AAGACCAAGCTGACCAATGAGAAACCCGCTGTGTACTAATAAAATACAAA--TAGCG 136276
1348 GGGGTGTGTAGCAGGCACTGTGTATCCAGCTACTTGGGAGGCTGAGGCAAGAAATCGCT 1407
136275 GGGGTGTGTGTGATCTGTGTATCCAGCTACTTGGGAGGCTGAGGCAAGAAATCGCT 136216
1408 TGAACCCAGAGAGTGGAGGTTGAGTGAAGTGAAGTGGCGGCTGCACTCCAGGCTGAG 1467
136215 TGAACCTGGAGAGTGGAGGTTGAGTGAAGTGAAGTGGCGGCTGCACTCCAGGCTGAG 136156
1468 CGACAGAGCAAGACTGTGTCTCAGAAAAAAGAAAAAAGAGAGAGAGAGAGAGAG 1527
136155 CAACAGAGATGAACTCCGTCTCAGAAAAAAGAAAAAAGAGAGAGAGAGAGAGAG 136096
1528 A 1528
136095 A 136095

RESULT 10
HS108F13      123101 bp      DNA      linear      PRI 21-MAR-2003
LOCUS      Homo sapiens chromosome X BAC CEPH108F13 map Xq13 markers
DEFINITION      PHKX1-DXS227, complete sequence.
ACCESSION      BX295542 AJ239320 AJ239324
VERSION      BX295542.1 GI:29150139
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 Ramser, J., Klages, S., Hemmig, S., Klein, M., Sudbrak, R., McDonnell, N.,
Chelly, J., Francis, F., Buckle, V., Reinhardt, R. and Lehrach, H.
Unpublished
2 (bases 1 to 123101)
MPING.
Direct Submission
Submitted (15-FEB-1999) MPING, Abt. Lehrach, Max Planck Institut
Fuer Molekulare Genetik, Innesstrasse 73, Berlin, 14195 Germany
config 01
1..123101.
Location/Qualifiers
1..123101
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"

```

```

/map="Xq13"
/clone="BAC CEPH108F13"
/clone_lib="CEPHB: human BAC library, creator: Denis Le
Paslier"
/note="region between markers PHKX1-DXS227"

ORIGIN

Query Match
Best Local Similarity 78.1%; Pred. No. 4,5e-49;
Matches 381; Conservative 0; Mismatches 104; Indels 3; Gaps 1;

14.1%; Score 305.6; DB 9; Length 123101;

1021 TAACTATATAGAGAGTTGTGACAAAACAGATGATTAAGCTCCGAACCTGGCACAAGC 1080
48208 TGAATATATATATATAGTGTGTCATATATATATATATATATATATATATATATAT 48267
1081 TCATAGTTCTAGCTGTGGAGGTTGAGGAGGAGGAGATGAGCTGTAACAGAGTTCAA 1140
48268 CTATATCTTAGCACTGTGGAGGCTGAGGAGGAGGAGATCCTGAGGCGGAGTTTCA 48327
1141 GGCAGCCTGGGCAACATTAACAGATCTGTCTCTCAAAAAAAGAAAAAAGAAAAA 1200
48328 GACCAAGCTGGCAACATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 48387
1201 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
48388 CAGAGTGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 48447
1261 GGGCGGATCACTGTGTGTGAGAGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
48448 AGGTAGGTCACTGTAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 48507
1321 CTGTACTCAAAATGCAAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
48508 CTCTACTTAAAGTAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 48567
1381 TTGGAGAGCTGA--GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
48568 TCGGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 48627
1438 TGAAGATCGTCCGTTGCACTCCAGCTGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1497
48628 TGAAGATGACCACTTACCTCCAGCTGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 48687
1498 AAAAAAA 1505
48688 AAAAAAA 48695

RESULT 11
AC084024/c      135404 bp      DNA      linear      PRI 19-SEP-2002
LOCUS      Homo sapiens chromosome 8, clone RP11-90P5, complete sequence.
DEFINITION      AC084024
ACCESSION      AC084024
VERSION      AC084024.17 GI:23196607
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 Birren, B., Nusbaum, C. and Lander, E.
Unpublished
2 (bases 1 to 135404)
RP11-90P5
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Boukhalil, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
Choquel, Y., Colangelo, M., Collins, S., Collins, S., Cooke, P.,
Deaile, M., Dewar, K., Diaz, J.S., Dodge, S., Fereira, P.,
Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,
Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,
Jilley, I., Johnson, R., Jones, C., Kann, L., Karatas, A., LaRoque, K.,
Lamaze, R., Lander, T., Lenczky, J., Levine, R., Liu, C., Liu, G.,

```



```
source
1. .195623
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8q23"
/clone="KB1198A1"
/cell_line="FLEB 14 - 14"
/cell_type="pre-pro-B cell"
/clone_lib="Keio BAC library"
376. .429
repeat_region
/evidence=not_experimental
/rpt_family="GA-rich"
1725. .1748
repeat_region
/evidence=not_experimental
/rpt_family="AT-rich"
1749. .2040
repeat_region
/evidence=not_experimental
/rpt_family="AluSx"
2041. .2071
repeat_region
/evidence=not_experimental
/rpt_family="AT-rich"
2261. .2322
repeat_region
/evidence=not_experimental
/rpt_family="L1MC5"
complement(2323. .2613)
/evidence=not_experimental
/rpt_family="AluDo"
2614. .2740
repeat_region
/evidence=not_experimental
/rpt_family="L1MC5"
2741. .3036
repeat_region
/evidence=not_experimental
/rpt_family="AluSx"
3040. .3330
repeat_region
/evidence=not_experimental
/rpt_family="AluY"
3331. .3350
repeat_region
/evidence=not_experimental
/rpt_family="L1MC5"
3351. .3653
repeat_region
/evidence=not_experimental
/rpt_family="AluDb"
3654. .3738
repeat_region
/evidence=not_experimental
/rpt_family="L1MC5"
complement(3739. .4244)
/evidence=not_experimental
/rpt_family="MLT1D"
4245. .4630
repeat_region
/evidence=not_experimental
/rpt_family="AluSp"
4631. .4716
repeat_region
/evidence=not_experimental
/rpt_family="AT-rich"
complement(4716. .5549)
/evidence=not_experimental
/rpt_family="L1MC4"
5549. .5622
repeat_region
/evidence=not_experimental
/rpt_family="TTRTA)n"
complement(5625. .5713)
/evidence=not_experimental
/rpt_family="L1MC4"
5942. .5996
repeat_region
/evidence=not_experimental
/rpt_family="AT-rich"
6056. .6339
repeat_region
/evidence=not_experimental
/rpt_family="AluSx"
6340. .6375
repeat_region
/evidence=not_experimental
/rpt_family="TAA)n"

repeat_region 6395. .6448
/evidence=not_experimental
/rpt_family="Alu"
6449. .6759
repeat_region
/evidence=not_experimental
/rpt_family="L1MB7"
6760. .7085
repeat_region
/evidence=not_experimental
/rpt_family="AluDo"
7086. .7242
repeat_region
/evidence=not_experimental
/rpt_family="L1MB7"
7243. .7707
repeat_region
/evidence=not_experimental
/rpt_family="MLT1D"
7708. .7736
repeat_region
/evidence=not_experimental
/rpt_family="L1MB7"
7741. .7862
repeat_region
/evidence=not_experimental
/rpt_family="L1MB7"
complement(7863. .8276)
/evidence=not_experimental
/rpt_family="MSTB"
8277. .8294
repeat_region
/evidence=not_experimental
/rpt_family="L1MB7"
8473. .9355
repeat_region
/evidence=not_experimental
/rpt_family="L1MB5"
complement(9370. .9867)
/evidence=not_experimental
/rpt_family="AluSx"
complement(10103. .10388)
/evidence=not_experimental
/rpt_family="AluDb"
10799. .11558
repeat_region
/evidence=not_experimental
/rpt_family="L1TR3g"
11590. .11704
repeat_region
/evidence=not_experimental
/rpt_family="AluDo"
11705. .11998
repeat_region
/evidence=not_experimental
/rpt_family="AluY"
11999. .12139
repeat_region
/evidence=not_experimental
/rpt_family="AluDo"
12173. .12248
repeat_region
/evidence=not_experimental
/rpt_family="AluDo/FlAM"
12262. .12299
repeat_region
/evidence=not_experimental
/rpt_family="AT-rich"
12371. .12680
repeat_region
/evidence=not_experimental
/rpt_family="TMB1B"
12808. .12949
repeat_region
/evidence=not_experimental
/rpt_family="I2"
13337. .13476
repeat_region
/evidence=not_experimental
/rpt_family="AluSx"
13477. .13768
repeat_region
/evidence=not_experimental
/rpt_family="AluSp"
13769. .13944
repeat_region
/evidence=not_experimental
/rpt_family="AluSx"
14073. .14197
repeat_region
/evidence=not_experimental
/rpt_family="I2"
14594. .14731
repeat_region
```



```

*      1      1787: contig of 1787 bp in length
*      1788      1887: gap of 100 bp
*      1888      15117: contig of 13230 bp in length
*      15118      15217: gap of 100 bp
*      15218      37920: contig of 22703 bp in length
*      37921      38020: gap of 100 bp
*      38021      79049: contig of 41029 bp in length
*      79050      79149: gap of 100 bp
*      79150      131081: contig of 51932 bp in length
*      131082      131181: gap of 100 bp
*      131182      172884: contig of 41703 bp in length.
FEATURES
SOURCE
1..172884
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-266D22"
/clone_1fb="RPC1-11 Human Male BAC"
misc_feature
1..1787
/clone="assembly_fragment"
clone_end:SP6
vector_side:left"
misc_feature
1888..15117
/clone="assembly_fragment"
15218..37920
/clone="assembly_fragment"
misc_feature
38021..79049
/clone="assembly_fragment"
79150..131081
/clone="assembly_fragment"
131182..172884
/clone="assembly_fragment"
misc_feature
172884..172884
/clone="assembly_fragment"
vector_side:right"
ORIGIN
Query Match      14.0%; Score 303.6; DB 2; Length 172884;
Best Local Similarity 77.1%; Pred. No. 11e-48;
Matches 387; Conservative 0; Mismatches 104; Indels 11; Gaps 1;
QY      1060 GCTGGCAACCGTGGACACGCTCATAGTCTTGAAGTCTTGGAGGTTGAGAGGAGAGAT 1119
DB      60049 GCTGGCGCCAGTGGTCACTCACTGTAATCTGACGACTTTGGAGGCTGAGTGGAGAT 59950
QY      1120 GCGTTGAACACAGTGTTCAGGCGCCTGGGCAACATACAGATCTGTCTCTCAA 1179
DB      59989 CACTTGAGCTCAGAGTTGAGACAGCCTGGGCAATAGAGAGAGCCCATCTCTATT 59930
QY      1180 AAAAAAAAA-----AAAAAAAAAGAGAGGCGCGGCGCTGTGCTCAGC 1228
DB      59929 TTTAATAAATAAATAATTTTAAAAAATAAATAAAGGCGCAGTGCCTCAAG 59870
QY      1229 CCTGTAATCCAGACCTTTGGAGGCGGCGGCGGATCACCTGTGTCAGAGATTG 1288
DB      59869 CCTGTAATCCAGACCTTTGGAGGCGGCGGCGGATCACCTGTGTCAGAGATTG 59810
QY      1289 AGACCAAGCTGGCCCAACATGCAAAACCCCGTCTGTACTCAAAATGCAAAATTAGCCAG 1348
DB      59809 AAGACCAAGCTGGCCCAACATGCAAAACCCCGTCTGTACTCAAAATGCAAAATTAGCTG 59750
QY      1349 GCGTGTAGCAGGACCTGTAAATCCAGTACTTGGAGGCTGAGGCGAGGATCGCTT 1408
DB      59749 GTGTGTAGCAGGCGCTGTAAATCCAGTACTTGGAGGCGGAGGCGAGGAAATCGCTT 59690
QY      1409 GAACCCAGAGGTTGAGGTTGAGTGAAGTGAAGTGCAGTGCCTGACATCCAGCTGGGC 1468
DB      59689 GAACCTGGAAGGTTGAGGTTGAGTGAAGTGAAGTGCAGTGCCTGACATCCAGCTGGGC 59630
QY      1469 GACCAAGAGCAACTGTCTCAGAAAAAATAAATAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1528
DB      59629 GACCAAGAGCAACTGTCTCAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 59570

```

```

QY      1529 ACAATATTGGAGAGAGAGAT 1550
DB      59569 TTAAGTTCTCCAGATAAGAT 59548
RESULT 15
AC024035
LOCUS
DEFINITION
AC024035
AC024035.3 GI:7637351
VERSION
AC024035
KEYWORDS
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 176063)
Waterston,R.H.
REFERENCE
1 (bases 1 to 176063)
Waterston,R.H.
AUTHORS
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 176063)
Waterston,R.H.
Direct Submission
Submitted (20-FEB-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Apr 22, 2000 this sequence version replaced gi:7024222.
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Project Information
Center project name: H_NH0266D22
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing method: Dye-terminator Big Dye; 0% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 170547 bases at least Q40
Consensus quality: 171287 bases at least Q30
Consensus quality: 171861 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 175063; sum-of-coverage
Quality coverage: 6.16 in Q20 bases; sum-of-coverage
Quality coverage: 6.14 in Q20 bases; sum-of-coverage
----- NOTE: This is a 'working draft' sequence. It currently
consists of 11 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1
2493: contig of 2493 bp in length
2494: gap of unknown length
2594: gap of 5355 bp in length
7948: gap of unknown length
7949: gap of unknown length
8048: gap of unknown length
8049: gap of 4998 bp in length
13047: gap of unknown length
13146: gap of unknown length
13147: gap of 6756 bp in length
19902: gap of unknown length
19903: gap of unknown length
20003: gap of 7777 bp in length
27779: contig of 7777 bp in length
27780: gap of unknown length
27879: gap of unknown length
27880: gap of 8800 bp in length
36679: gap of unknown length
36680: gap of unknown length
36780: gap of 15557 bp in length
52037: gap of unknown length
52136: gap of unknown length
71975: contig of 19839 bp in length

```

FEATURES
* 71976 72075: gap of unknown length
* 72076 97520: contig of 25445 bp in length
* 97521 97620: gap of unknown length
* 97621 133840: contig of 36220 bp in length
* 133841 133940: gap of unknown length
* 133941 176063: contig of 42123 bp in length.
Location/Qualifiers
1. 176063
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/clone="RP11-266D22"
1. 2493
/note="assembly_name:Contig23"
2594. 7948
/note="assembly_name:Contig24"
8049. 13046
/note="assembly_name:Contig25
clone_end:SP6
vector_side:left"
13147. 19902
/note="assembly_name:Contig26"
20003. 27779
/note="assembly_name:Contig27"
27880. 36679
/note="assembly_name:Contig28"
36780. 52036
/note="assembly_name:Contig29"
52137. 71975
/note="assembly_name:Contig30"
72076. 97520
/note="assembly_name:Contig31"
97621. 133840
/note="assembly_name:Contig32"
133941. 176063
/note="assembly_name:Contig33
clone_end:T7
vector_side:left"

misc_feature
/note="assembly_name:Contig23"
2594. 7948
/note="assembly_name:Contig24"
8049. 13046
/note="assembly_name:Contig25
clone_end:SP6
vector_side:left"
13147. 19902
/note="assembly_name:Contig26"
20003. 27779
/note="assembly_name:Contig27"
27880. 36679
/note="assembly_name:Contig28"
36780. 52036
/note="assembly_name:Contig29"
52137. 71975
/note="assembly_name:Contig30"
72076. 97520
/note="assembly_name:Contig31"
97621. 133840
/note="assembly_name:Contig32"
133941. 176063
/note="assembly_name:Contig33
clone_end:T7
vector_side:left"

ORIGIN

Query Match 14.0%; Score 303.6; DB 2; Length 176063;
Best Local Similarity 77.1%; Pred. No. 1.1e-48;
Matches 387; Conservative 0; Mismatches 104; Indels 11; Gaps 1;

QY 1060 GCTGCAACCGTGGCAACGCTCATAGTCTAGCTCTTGGAGGTTGAGAGGAGGAT 1119
DB 30505 GCTGGGCGCAGTGGCTCACACTGTATCTCAGCACTTTGGAGGCTGAGGTGGAGGAT 30564
QY 1120 GGCTTGAACACAGGTGTTCAAGGCCAGCTGGGCAACATACAAAGTCTGTCTCTCAAA 1179
DB 30565 CACTTGAAGCTCAGAGTTCAGACCAAGCTGGGCAATAGAGAGCCCATCTCTATT 30624
QY 1180 AAAAAAAAA-----AAAAAAAAAGAGAGAGGCGCGGCGTGTGGCTCAG 1228
DB 30625 TTTAAAAAATAATATTTTAAAAAATAAATAAAGGCGCAGGTGCGTGGCTCAG 30684
QY 1229 CCTGTATCCACACACTTTGGAGGCGGAGCCGCGGATCACTTGTGTGAGATTG 1288
DB 30685 CCTGTATCCACACACTTTGGAGGCGTGAAGGAGGAGATCACTTGTGTGAGATTG 30744
QY 1289 AGACCAAGCTGGGCAACATGGCAAAACCCGCTGTACTCAAAATGCAAAATTAGCCAG 1348
DB 30745 AGACCAAGCTGGGCAACATGGTAAACCCGCTGTACTCAAAATGCAAAATTAGCTGG 30804
QY 1349 GCGGTGTAAGAGGCACTGTAAATCCAGGTAATTGGAGGCTGAGGCAAGATCGCTT 1408
DB 30805 GTGTGTGGAGGCGGCTGTAAATCCAGGTAATTGGAGGCTGAGGCAAGATCGCTT 30864
QY 1409 GAACCGAGAGGTGAGGTTGACGTAAAGTGAATGTCGCTTGCCTCCAGCTTGGGC 1468
DB 30865 GAACGTGGAGGTGAGGTTGACGTAAAGTGAATGTCGCTTGCCTCCAGCTTGGGC 30924

QY 1469 GACAAAGCAAGCTCTGTCTCAGAAAAAAGAGAGAGAGAGAAAGAGA 1528
DB 30925 GACAAAGCAAGACTTCACTCAAAAAATAAATAAATAAATCAAGCATCA 30984
QY 1529 ACAATATTGGAGAGAAAGAT 1550
DB 30985 TTAAGTTCTCCAGATAAGAT 31006

Search completed: March 12, 2004, 21:22:39
Job time : 6846.04 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 12:26:39 ; Search time 869.74 Seconds

(without alignments)
10609.014 Million cell updates/sec

Title: US-09-966-880A-15

Perfect score: 2172
Sequence: 1 cccctgataggtgctgata.....agatgataggataaactgt 2172

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2002s:*
6: geneseqn2003s:*
7: geneseqn2004s:*
8: geneseqn2005s:*
9: geneseqn2006s:*
10: geneseqn2007s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 2172 | 100.0 | 2172 | 3 | AAC55319 Human act |
| 2 | 2172 | 100.0 | 2791 | 6 | AB573287 DNA encod |
| 3 | 2172 | 100.0 | 2791 | 6 | AB573288 DNA encod |
| 4 | 2172 | 100.0 | 2818 | 3 | AAC55312 Human act |
| 5 | 2172 | 100.0 | 6564 | 3 | AAC55314 Human act |
| 6 | 2172 | 100.0 | 11204 | 3 | AAC55339 Human act |
| 7 | 2172 | 100.0 | 11204 | 6 | AB573286 DNA encod |
| 8 | 1488 | 68.5 | 1543 | 7 | ABX05468 Human nov |
| 9 | 923 | 42.5 | 1543 | 7 | ABX05468 Human nov |
| 10 | 574 | 26.4 | 574 | 4 | AAK81089 Human imm |
| 11 | 429 | 19.8 | 429 | 4 | AAK61819 Human imm |
| 12 | 295.2 | 13.6 | 13862 | 4 | AAI02789 Human rep |
| 13 | 295.2 | 13.6 | 13862 | 4 | AAI07516 Human rep |
| 14 | 295.2 | 13.6 | 13862 | 4 | ABX08208 Human cva |
| 15 | 293.8 | 13.5 | 110000 | 7 | ACP42745_0 Human ALM |
| 16 | 292.6 | 13.5 | 7809 | 5 | AAI03406 Human rep |
| 17 | 292.6 | 13.5 | 7809 | 5 | AB15724 Human ner |
| 18 | 292.4 | 13.5 | 30393 | 4 | AAK67339 Human imm |
| 19 | 291 | 13.4 | 25012 | 5 | AB15531 Human ner |
| 20 | 291 | 13.4 | 25012 | 7 | AB73798 Secreted |
| 21 | 291 | 13.4 | 25012 | 7 | AD98470 Human sec |
| 22 | 288.8 | 13.3 | 96596 | 8 | ADN02864 Human TTP |
| 23 | 288.8 | 13.3 | 96596 | 9 | ADB72602 Human ITP |

| | | | | | | | | |
|---|----|-------|------|--------|---|-----------|-----------|-----------|
| C | 24 | 288.8 | 13.3 | 96596 | 9 | AD85343 | AD85343 | Mouse ITP |
| C | 25 | 287.8 | 13.3 | 110000 | 9 | AD81169_0 | AD81169_0 | Human tra |
| C | 26 | 286.2 | 13.2 | 37314 | 4 | AAK71358 | AAK71358 | Human tra |
| C | 27 | 286.2 | 13.2 | 172570 | 6 | AB088207 | AB088207 | Human tra |
| C | 28 | 284.4 | 13.1 | 32249 | 4 | AAI04676 | AAI04676 | Human rep |
| C | 29 | 284.4 | 13.1 | 32249 | 4 | AB197583 | AB197583 | Human tes |
| C | 30 | 284.4 | 13.1 | 129722 | 6 | AB088117 | AB088117 | Human tra |
| C | 31 | 283.4 | 13.0 | 1792 | 6 | ABX96614 | ABX96614 | Human tra |
| C | 32 | 283.2 | 13.0 | 22428 | 4 | AA541759 | AA541759 | Genomic s |
| C | 33 | 283.2 | 13.0 | 22428 | 7 | AB267767 | AB267767 | Human sec |
| C | 34 | 283.2 | 13.0 | 22428 | 7 | AB74201 | AB74201 | Secreted |
| C | 35 | 283.2 | 13.0 | 22428 | 7 | ADA98730 | ADA98730 | Human sec |
| C | 36 | 282.4 | 13.0 | 13919 | 6 | ABX86218 | ABX86218 | DNA encod |
| C | 37 | 282.4 | 13.0 | 13919 | 6 | ABX86221 | ABX86221 | AIP-1/FLA |
| C | 38 | 282.4 | 13.0 | 13919 | 6 | ABX86220 | ABX86220 | AIP-1/FLA |
| C | 39 | 282.4 | 13.0 | 21404 | 6 | ABX86229 | ABX86229 | AIP-1/FLA |
| C | 40 | 281.4 | 13.0 | 65608 | 6 | AB162910 | AB162910 | Breast ca |
| C | 41 | 281.4 | 13.0 | 65608 | 6 | AB164414 | AB164414 | Stomach c |
| C | 42 | 281.4 | 13.0 | 65608 | 6 | AB167668 | AB167668 | Oesophagu |
| C | 43 | 281 | 12.9 | 23452 | 4 | AA542122 | AA542122 | Genomic s |
| C | 44 | 280.2 | 12.9 | 11821 | 4 | AA136492 | AA136492 | Human mus |
| C | 45 | 280.2 | 12.9 | 11821 | 7 | ABX59480 | ABX59480 | CDNA enco |

ALIGNMENTS

| | | |
|----------|--|-------------------------|
| RESULT 1 | AAC55319 | standard; DNA; 2172 BP. |
| ID | AAC55319 | |
| AC | AAC55319; | |
| DT | 05-FEB-2001 | (first entry) |
| XX | Human activation-induced cytidine deaminase exon 5 SEQ ID NO:15. | |
| XX | Activation-induced cytidine deaminase; AID; cytidine deaminase; immune related disease; allergy; allergic disease; antileukemic; antineoplastic; ophthalmological; anti-HIV; dermatological; gene therapy; B cell associated immune system disorder; food allergy; immunodeficiency disease; immunoglobulin A deficiency disease; asthma; IGA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis; drug allergy; allergic rhinitis; Rosen disease; Digeorge disease; AIDS; ataxia telangiectasia; common variable immunodeficiency disorder; major histocompatibility class II deficiency disease; auto immunodeficiency syndrome; IgG subclass selection disorder; ds. | |
| XX | Homo sapiens. | |
| XX | WO200058480-A1. | |
| XX | 05-OCT-2000. | |
| PD | 28-MAR-2000; 2000MO-JP001918. | |
| XX | 29-MAR-1999; 99JP-00087192. | |
| PR | 24-JUN-1999; 99JP-00178999. | |
| PR | 27-DEC-1999; 99JP-00371382. | |
| XX | (NISR) JAPAN TOBACCO INC. | |
| PA | (HONU) HONU T. | |
| PI | Honjo T, Muramatsu M; | |
| XX | WPI; 2000-611715/58. | |
| DR | Nucleic acid encoding activation induced cytidine deaminase, useful as a target for drug development for immune-related diseases including allergies. | |
| PT | Claim 16; Page 152-153; 174pp; Japanese. | |
| XX | | |

DB 1801 AGAGTAAATGTTACTGATGATTCATTATATTTATTTATTTGGCTCTATG 1860
 QY 1861 ATTTTATTAACATGATTTCTTTCTGATATATGAAAGAGTCAAGCTTATA 1920
 DB 1861 ATTTTATTAACATGATTTCTTTCTGATATATGAAAGAGTCAAGCTTATA 1920
 QY 1921 AATTTTACTTTAGAAATGATTCATTAACACAGATATGTAATTCATTCAGTAT 1980
 DB 1921 AATTTTACTTTAGAAATGATTCATTAACACAGATATGTAATTCATTCAGTAT 1980
 QY 1981 GATGCTACGAGACCATTTCTCTGATTTTATGTAATTCATTCAGTATTCGCTC 2040
 DB 1981 GATGCTACGAGACCATTTCTCTGATTTTATGTAATTCATTCAGTATTCGCTC 2040
 QY 2041 TGGCTCATCTTCAATCAGTATTAATTAATTAATTTTGAAGCTGTGAGATATAA 2100
 DB 2041 TGGCTCATCTTCAATCAGTATTAATTAATTAATTTTGAAGCTGTGAGATATAA 2100
 QY 2101 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160
 DB 2101 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160
 QY 2161 GGAATTAATCTTG 2172
 DB 2161 GGAATTAATCTTG 2172
 RESULT 2
 ABS73287
 ID ABS73287 standard; DNA; 2791 BP.
 XX
 AC ABS73287;
 XX
 DT 04-DEC-2002 (first entry)
 XX
 DE DNA encoding human translocation del(12p) protein #2.
 XX
 KM Chromosome aberration; oncogenic fusion protein; cancer;
 KM proliferative disease; cellular fusion isoform; heat shock protein 90;
 KM HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
 KM T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
 KM acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
 KM acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
 KM papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
 KM rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.
 OS Homo sapiens.
 XX
 PN WO200269900-A2.
 XX
 PD 12-SEP-2002.
 XX
 PF 01-MAR-2002; 2002WO-US006518.
 XX
 PR 01-MAR-2001; 2001US-0272751P.
 XX
 PA (CONF-) CONFORMA THERAPEUTICS CORP.
 XX
 PI Filtz LC, Burrows FU;
 XX
 DR WPI: 2002-699710/75.
 DR P-PSDB; ABG95083.
 XX
 PT Treating genetically-defined disease associated with chromosomal
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
 PT diseases, involves administering an inhibitor of heat shock protein 90.
 XX
 PS Disclosure; Page 246-247; 389pp; English.
 CC The invention describes a method of treating genetically-defined disease
 CC associated with chromosomal aberrations yielding oncogenic fusion
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous
 CC cell population, treating proliferative diseases associated with mutant

CC protein or cellular protein isoforms (II) dependent on heat shock protein
 CC (HSP)-90, or selectively treating cells expressing (II) involving
 CC administering HSP90-inhibitor. The method is useful for treating
 CC genetically-defined disease with chromosomal aberration yielding
 CC oncogenic fusion protein, treating cancerous cells containing fusion
 CC protein in heterogeneous cell population, treating proliferative disease
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
 CC p53), or selectively treating cells expressing mutant protein or cellular
 CC protein isoform in a patient heterozygous for (II). The method is useful
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CMML,
 CC or a disease characterised by a solid tumour such as papillary thyroid
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
 CC synovial sarcoma. The method is also useful for treating viral
 CC infections. This represents the DNA sequence of a chromosome aberration
 XX
 SQ Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;
 Query Match 100.0%; Score 2172; DB 6; Length 2791;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCCCTGTATGAGCTTGATGATCTTACGAGACGATTTGCTACTTTGGACCTTGATAGCA 60
 DB 620 CCCCTGTATGAGCTTGATGATCTTACGAGACGATTTGCTACTTTGGACCTTGATAGCA 679
 QY 61 CTTCCAGGAATGTCACACACGATGAAATATCTCTGCTGAAGACATGATATAAATACGT 120
 DB 680 CTTCCAGGAATGTCACACACGATGAAATATCTCTGCTGAAGACATGATATAAATACGT 739
 QY 121 CCTCAGAGCTCTCTGTTTATTTTCTTCACTCTCATCTTCTTGAAGTTTACAGAAAAA 180
 DB 740 CCTCAGAGCTCTCTGTTTATTTTCTTCACTCTCATCTTCTTGAAGTTTACAGAAAAA 799
 QY 181 AATTTATATACGACTCTTTTAAAAAGATCTATGCTTGAATAATAGAGAGAACACAGT 240
 DB 800 AATTTATATACGACTCTTTTAAAAAGATCTATGCTTGAATAATAGAGAGAACACAGT 859
 QY 241 CTGGCCAGGAGAGCTGCTGCAATTTGCTGAGTTTGAATGCAACATTTGCCCTACTGGA 300
 DB 860 CTGGCCAGGAGAGCTGCTGCAATTTGCTGAGTTTGAATGCAACATTTGCCCTACTGGA 919
 QY 301 ATTAACAGACCTGACAGACCTGGAGACCTTAAGTGTCAACGTTTCTTCAAGCTTTA 360
 DB 920 ATTAACAGACCTGACAGACCTGGAGACCTTAAGTGTCAACGTTTCTTCAAGCTTTA 979
 QY 361 GGTAGATGAGACAGAAAGTATCTTAAAAAGCATGGTGAAGATCAAAATGTTTTTA 420
 DB 980 GGTAGATGAGACAGAAAGTATCTTAAAAAGCATGGTGAAGATCAAAATGTTTTTA 1039
 QY 421 TATCAACATCTTTATATTTATTTGATTCATTTAAGTTTAAACAGTGTGATAGATTTT 480
 DB 1040 TATCAACATCTTTATATTTATTTGATTCATTTAAGTTTAAACAGTGTGATAGATTTT 1099
 QY 481 TCTATTTCTTTCCCTGAGCTTACTTTCAAGTAAACAAACTCTTCCATCGAGGCATGA 540
 DB 1100 TCTATTTCTTTCCCTGAGCTTACTTTCAAGTAAACAAACTCTTCCATCGAGGCATGA 1159
 QY 541 TCTATGAGACCTCTTAATGAGATCTGGGATGTTGAGACCCCAACATCTTCCAAA 600
 DB 1160 TCTATGAGACCTCTTAATGAGATCTGGGATGTTGAGACCCCAACATCTTCCAAA 1219
 QY 601 GCATTAATATCCATATGCGCTGTATGTTTAAATCAGACAGAGCATGTTTATGTTTG 660
 DB 1220 GCATTAATATCCATATGCGCTGTATGTTTAAATCAGACAGAGCATGTTTATGTTTG 1279
 QY 661 TACAAAAGAAATGTTATGAGTGGGATGAGAGTATAGACATGATGATCACTTCAA 720
 DB 1280 TACAAAAGAAATGTTATGAGTGGGATGAGAGTATAGACATGATGATCACTTCAA 1339
 QY 721 GCTACTTATATAGATCTTAAAAATGGGACAGAGCATGTGAACAAACACCTTAATA 780

| | | | |
|----|------|---|------|
| Db | 1340 | GCTACTTAAATAAGATCTTAAATGGGCGAGGAGACTGTGAACAAGACACCTATATA | 1399 |
| Qy | 781 | TGGGTGATGTCCTAGATGACAAATCTTCTGGAAACGCAAACTCTTTAAGAAATCCCT | 840 |
| Db | 1400 | TGGGTATATGCTCAAGTAGCAAAATCTCTGGAAACGCAAACTCTTTAAGAAATCCCT | 1459 |
| Qy | 841 | AATTAGAAACCCACCAACTTACATATATATATAGCAAACAATTGGAAGAAATTG | 900 |
| Db | 1460 | AATTTAGAACCCCAAACTTACATATATATATAGCAAACAATTGGAAGAAATTG | 1519 |
| Qy | 901 | CTTGAAATGTTGGGAGAGAAAATCTATGTGCTCTGTGGTCTCTTCATCTCAGAAATG | 960 |
| Db | 1520 | CTTGAAATGTTGGGAGAGAAAATCTATGTGCTCTGTGGTCTCTTCATCTCAGAAATG | 1579 |
| Qy | 961 | CCAAATCAGGTCAGAGTTTGTCTACATTTTGTATGTGTGATGCTTCTCCAAAGTATAT | 1020 |
| Db | 1580 | CCAAATCAGGTCAGAGTTTGTCTACATTTTGTATGTGTGATGCTTCTCCAAAGTATAT | 1639 |
| Qy | 1021 | TAACTATATAGAGATGTGTACAAACAGATATATAAGCTGTGGAAACCGGAGACAGC | 1080 |
| Db | 1640 | TAACTATATAGAGATGTGTACAAACAGATATATAAGCTGTGGAAACCGGAGACAGC | 1699 |
| Qy | 1081 | TCATATGTTCTAGTCTGCTGGAGAGTTGAGAGGAGATGCTTGAAACAAGTGTTCAA | 1140 |
| Db | 1700 | TCATATGTTCTAGTCTGCTGGAGAGTTGAGAGGAGATGCTTGAAACAAGTGTTCAA | 1759 |
| Qy | 1141 | GGCCAGCTGGGACACATMACAAGTCTGTCTCCAAAAAATAAAAAAAAAAAGAAA | 1200 |
| Db | 1760 | GGCCAGCTGGGACACATMACAAGTCTGTCTCCAAAAAATAAAAAAAAAAAGAAA | 1819 |
| Qy | 1201 | GAGAGAGGCGCGGCGTGTGGTCTCAGCCCTGTATCCCAACCTTTGGAGGCGGAGCC | 1260 |
| Db | 1820 | GAGAGAGGCGCGGCGTGTGGTCTCAGCCCTGTATCCCAACCTTTGGAGGCGGAGCC | 1879 |
| Qy | 1261 | GAGGCGGATCACCTGTGGTCAAGAGTTTACAACAAGCTGTGGCAACATGCAAAACCCGT | 1320 |
| Db | 1880 | GAGGCGGATCACCTGTGGTCAAGAGTTTACAACAAGCTGTGGCAACATGCAAAACCCGT | 1939 |
| Qy | 1321 | CTGTACTCMAAATGCAAAAATTAGCCAGGCGTGTAGGAGCACCCTGTATCCCAAGTAC | 1380 |
| Db | 1940 | CTGTACTCMAAATGCAAAAATTAGCCAGGCGTGTAGGAGCACCCTGTATCCCAAGTAC | 1999 |
| Qy | 1381 | TTGGAGGCTGAGGAGAGAGATGCTTGAACCAAGAGGTGTAGAGTTGCACTAAGCTGA | 1440 |
| Db | 2000 | TTGGAGGCTGAGGAGAGAGATGCTTGAACCAAGAGGTGTAGAGTTGCACTAAGCTGA | 2059 |
| Qy | 1441 | GATGCTGCGTTGCACTCCAGCTGTGGGAGCAAGAGCAAGACTGTCTCAGAAAAAAA | 1500 |
| Db | 2060 | GATGCTGCGTTGCACTCCAGCTGTGGGAGCAAGAGCAAGACTGTCTCTCAGAAAAAAA | 2119 |
| Qy | 1501 | AAAAAAAAAGAGAGAGAGAGAAAGACAATATTTGGAGAGAGAGTGGGAGAGAT | 1560 |
| Db | 2120 | AAAAAAAAAGAGAGAGAGAGAAAGACAATATTTGGAGAGAGAGTGGGAGAGAT | 2179 |
| Qy | 1561 | TGCAAGAAATTTGCTTTATTCACAACAATGTAGAAGCCATATAGGGATTCCTATTTG | 1620 |
| Db | 2180 | TGCAAGAAATTTGCTTTATTCACAACAATGTAGAAGCCATATAGGGATTCCTATTTG | 2239 |
| Qy | 1621 | TCTCTTTTGGTGTCTAATTTGTCCCTAACAACCTGTCTTTGACAGAGAAAAATTCGA | 1680 |
| Db | 2240 | TCTCTTTTGGTGTCTAATTTGTCCCTAACAACCTGTCTTTGACAGAGAAAAATTCGA | 2299 |
| Qy | 1681 | ATAACCATATCCCTGTGCGTATATATCTAGAACCTTGTGAATGAATAGAGAGATCC | 1740 |
| Db | 2300 | ATAACCATATCCCTGTGCGTATATATCTAGAACCTTGTGAATGAATAGAGAGATCC | 2359 |
| Qy | 1741 | ACAGGAAACTTGAATGACAACTGTCTTATTTAATCTTATGTGACATAAAGTTGTAA | 1800 |
| Db | 2360 | ACAGGAAACTTGAATGACAACTGTCTTATTTAATCTTATGTGACATAAAGTTGTAA | 2419 |
| Qy | 1801 | AGAGTTAAAAATTTTACTCTCACTGATTCATTTAATTTTATATTTTGGCTATAG | 1866 |
| Db | 2420 | AGAGTTAAAAATTTTACTCTCACTGATTCATTTAATTTTATATTTTGGCTATAG | 2479 |

| | | | |
|----------|---|--|-----|
| QY | 1861 | ATTATTATTAACATGATTTCTCTTCTGAAATATGAAATGGAGCTCAAAGCTTCATA | 192 |
| Db | 2480 | ATTTTATTAACAATGATTTCTCTTCTGAAATATGAAATGGAGCTCAAAGCTTCATA | 253 |
| QY | 1921 | AATTATTAACCTTAGAATGATTCGAATTAACAAGATGATGTAATGTGAATGCACTGATAT | 198 |
| Db | 2540 | AATTATTAACCTTAGAATGATTCGAATTAACAAGATGATGTAATGTGAATGCACTGATAT | 259 |
| QY | 1981 | GGGCTACGACGACCATTTCTCTGATTTTATGTAACCTTTATGACGCAATTTGGCTTC | 204 |
| Db | 2600 | GGGCTACGACGACCATTTCTCTGATTTTATGTAACCTTTATGACGCAATTTGGCTTC | 265 |
| QY | 2041 | TGGCTCACTTCAATCACTGATTAATTAATGATTAATATTTTGGAGCTGTGAAGATTAATA | 210 |
| Db | 2660 | TGGCTCACTTCAATCACTGATTAATTAATGATTAATATTTTGGAGCTGTGAAGATTAATA | 271 |
| QY | 2101 | TACCAATTAATTAATTAATTAAGATGTTTATTAATGATTAATTAATTAATTAATGATGAT | 216 |
| Db | 2720 | TACCAATTAATTAATTAATTAAGATGTTTATTAATGATTAATTAATTAATTAATGATGAT | 277 |
| QY | 2161 | GGATTAACCTG 2172 | |
| Db | 2780 | GGATTAACCTG 2791 | |
| RESULT 3 | | | |
| ID | ABS73288 | standard; DNA; 2791 BP. | |
| AC | ABS73288; | | |
| DT | 04-DEC-2002 | (first entry) | |
| DE | DNA encoding human translocation del(12p) protein #3. | | |
| KW | Chromosome aberration; oncogenic fusion protein; cancer; | | |
| KW | proliferative disease; cellular protein isoform; heat shock protein 90; | | |
| KW | HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder; | | |
| KW | T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML; | | |
| KW | acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML; | | |
| KW | acute lymphoblastic leukaemia; ALL; APV; NHL; solid tumour; | | |
| KW | papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma; | | |
| KW | rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds. | | |
| XX | Homo sapiens. | | |
| OS | | | |
| FN | WO200269900-A2. | | |
| XX | | | |
| XX | 12-SEP-2002. | | |
| XX | | | |
| XX | 01-MAR-2002; 2002WO-US006518. | | |
| XX | | | |
| XX | 01-MAR-2001; 2001US-027251P. | | |
| XX | | | |
| PA | (CONF-) CONFORMA THERAPEUTICS CORP. | | |
| XX | | | |
| XX | Fritz LC; Burrows FU; | | |
| XX | | | |
| XX | WPI; 2002-698710/75. | | |
| XX | | | |
| XX | P-PSDB; ABS95084. | | |
| XX | | | |
| XX | | | |
| XX | Disclosure; Page 248-249; 389pp; English. | | |
| CC | The invention describes a method of treating genetically-defined disease | | |
| CC | associated with chromosomal aberrations yielding oncogenic fusion | | |
| CC | proteins (I), treating cancerous cells containing (I) in a heterogeneous | | |
| CC | cell population, treating proliferative diseases associated with mutant | | |
| CC | protein or cellular protein isoforms (II) dependent on heat shock protein | | |

(HSP)-90, or selectively treating cells expressing (ii) involving administering HSP90-inhibitor. The method is useful for treating genetically-defined disease with chromosomal aberration yielding oncogenic fusion protein, treating cancerous cells containing fusion protein in heterogeneous cell population, treating proliferative disease (e.g. rheumatoid arthritis or cancer) associated with mutant protein or cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g. p53), or selectively treating cells expressing mutant protein or cellular protein isoform in a patient heterozygous for (ii). The method is useful for treating a disease e.g. haematopoietic disorder such as T or B cell lymphoma, chronic myeloid leukaemia (CML), APL, ALL, AML, NHL and CML, or a disease characterised by a solid tumour such as papillary thyroid carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and synovial sarcoma. The method is also useful for treating viral infections. This represents the DNA sequence of a chromosome aberration

XX Sequence 2791 BP, 842 A, 548 C, 625 G, 776 T, 0 U, 0 Other;

Query Match 100.0%; Score 2172; DB 6; Length 2791;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCCTGTATGAGTGTAGTACTTACGAGACGCAATTCGACTTTGGAGCTTGTATAGCA 60
620 CCCCTGTATGAGTGTAGTACTTACGAGACGCAATTCGACTTTGGAGCTTGTATAGCA 679

QY 61 CTTCAGAGATGTCAACACGATGAAATCTCTGCTGAAGACAGTGGATAAAAACAGT 120
DB 680 CTTCAGAGATGTCAACACGATGAAATCTCTGCTGAAGACAGTGGATAAAAACAGT 739

QY 121 CCTGAAGCTCTCTGTTTTTATCTTCACTCCTCTTCTTGAATTTACAGAAAA 180
DB 740 CCTGAAGCTCTCTGTTTTTATCTTCACTCCTCTTCTTGAATTTACAGAAAA 799

QY 181 ATATTTATATACGACTCTTTAAAAAGATCTATGTCTTGAATAAGAGAAACAGAGT 240
DB 800 ATATTTATATACGACTCTTTAAAAAGATCTATGTCTTGAATAAGAGAAACAGAGT 859

QY 241 CTGGCCAGGAGCGTCTGCAATTTGTCAGTTTGAATGCAACATTTGCCCTACTGGA 300
DB 860 CTGGCCAGGAGCGTCTGCAATTTGTCAGTTTGAATGCAACATTTGCCCTACTGGA 919

QY 301 ATAAAGAGACTCGAGAGCTGGAGACATCTTAAAGTGTCAAGCTTTCTATGACTTTA 360
DB 920 ATAAAGAGACTCGAGAGCTGGAGACATCTTAAAGTGTCAAGCTTTCTATGACTTTA 979

QY 361 GGTAGAGTAGAGCAGAGAGTAGATCTTAAAAAGATGGTAGAGAGATCAATGTTTTTA 420
DB 980 GGTAGAGTAGAGCAGAGAGTAGATCTTAAAAAGATGGTAGAGAGATCAATGTTTTTA 1039

QY 421 TATCAACATCTTTTATTTATTTGATTCATTGATTAACAGTGGTGTATAGATATTTT 480
DB 1040 TATCAACATCTTTTATTTATTTGATTCATTGATTAACAGTGGTGTATAGATATTTT 1099

QY 481 TCTATTTCTTTTCTTGTACGTTTACTTTCAGTAACAACAACCTTTCATCAGGCCATGA 540
DB 1100 TCTATTTCTTTTCTTGTACGTTTACTTTCAGTAACAACAACCTTTCATCAGGCCATGA 1159

QY 541 TCTATAGAGCTCTTAAAGAGTATCTGGGTGATTTGACCCCAACCACTCTCCCAA 600
DB 1160 TCTATAGAGCTCTTAAAGAGTATCTGGGTGATTTGACCCCAACCACTCTCCCAA 1219

QY 601 GCATTAATATCCATCATGCGCTGATGTTTAAATCAGAGAGAGATGTTTATGTTTG 660
DB 1220 GCATTAATATCCATCATGCGCTGATGTTTAAATCAGAGAGAGATGTTTATGTTTG 1279

QY 661 TACAAAAGAGATGTTATGGGTGGAGTAGAGGTATAGACCATGCTGACACTTCAA 720
DB 1280 TACAAAAGAGATGTTATGGGTGGAGTAGAGGTATAGACCATGCTGACACTTCAA 1339

QY 721 GCTACTTTATATAAGATCTTAAAAATGGCAGAGAGACTGTGACAAAGACCTTAATA 780
DB 1340 GCTACTTTATATAAGATCTTAAAAATGGCAGAGAGACTGTGACAAAGACCTTAATA 1399

QY 781 TGGGTGATGTCTGAGTAGCAATCTTCTGGAAGCAAACTCTTTAAGAGAGTCCCT 840
DB 1400 TGGGTGATGTCTGAGTAGCAATCTTCTGGAAGCAAACTCTTTAAGAGAGTCCCT 1459

QY 841 AATTGAAAACACCAACCACTTCAATATATTTAGCAACAAATTTGAAAGAGTTG 900
DB 1460 AATTGAAAACACCAACCACTTCAATATATTTAGCAACAAATTTGAAAGAGTTG 1519

QY 901 CTGGAATGTGGGAGAGAGAAAAATCTATTTGCTCTGTGGTCTCTTCACTTCAGAAAAG 960
DB 1520 CTGGAATGTGGGAGAGAGAAAAATCTATTTGCTCTGTGGTCTCTTCACTTCAGAAAAG 1579

QY 961 CCAATCAGGTCAGGTTTGTACATTTTATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1580 CCAATCAGGTCAGGTTTGTACATTTTATGATGATGATGATGATGATGATGATGATGAT 1639

QY 1021 TAACTATATAGAGAGTGTGACAAAACAGAAATGATTAAGCTGCAACCGTGGACACGC 1080
DB 1640 TAACTATATAGAGAGTGTGACAAAACAGAAATGATTAAGCTGCAACCGTGGACACGC 1699

QY 1081 TCATAGTTCTAGCTGCTTGGAGGTTGAGAGAGGAGATGCTTGAACACAGGTCTCA 1140
DB 1700 TCATAGTTCTAGCTGCTTGGAGGTTGAGAGAGGAGATGCTTGAACACAGGTCTCA 1759

QY 1141 GGCACGCTGGGACAAATACAAAGATCTCTCTCAAAAAAAAAAAAAAAAAAGAA 1200
DB 1760 GGCACGCTGGGACAAATACAAAGATCTCTCTCAAAAAAAAAAAAAAAAAAGAA 1819

QY 1201 GAGAGAGGCGCGGCGTGGTGGCTCAGCGCTGTAATCCAGACCTTTGGAGAGCGAGCC 1260
DB 1820 GAGAGAGGCGCGGCGTGGTGGCTCAGCGCTGTAATCCAGACCTTTGGAGAGCGAGCC 1879

QY 1261 GGGCGGATCACTGTGTCAAGAGTTGAGACCAAGCTTGGCAATGGCAAAACCCCGT 1320
DB 1880 GGGCGGATCACTGTGTCAAGAGTTGAGACCAAGCTTGGCAATGGCAAAACCCCGT 1939

QY 1321 CTGTACTCAAAATGCAAAATTTAGCGCGGTGTGACAGGACCTGTATCTCCAGCTTC 1380
DB 1940 CTGTACTCAAAATGCAAAATTTAGCGCGGTGTGACAGGACCTGTATCTCCAGCTTC 1999

QY 1381 TTGGAGGCTGAGGCGAGAGAAATCGCTGAACCCAGAGAGTGGAGGTTGACATTAAGCTGA 1440
DB 2000 TTGGAGGCTGAGGCGAGAGAAATCGCTGAACCCAGAGAGTGGAGGTTGACATTAAGCTGA 2059

QY 1441 GATCGTGCCTTGCATCTCAGCTGGGCGCAAGAGCAAGACTCTGTCTCAAAAAAAAA 1500
DB 2060 GATCGTGCCTTGCATCTCAGCTGGGCGCAAGAGCAAGACTCTGTCTCAAAAAAAAA 2119

QY 1501 AAAAAAAGAGAGAGAGAGAAAGAAACAATTTTGGAGAGAGAGATGGGAGAGCAT 1560
DB 2120 AAAAAAAGAGAGAGAGAGAGAAAGAAACAATTTTGGAGAGAGAGATGGGAGAGCAT 2179

QY 1561 TGCAGAGAAATGTGCTTTATGCAAAATGTAAGAGACCAATTAAGGATCCCTATTG 1620
DB 2180 TGCAGAGAAATGTGCTTTATGCAAAATGTAAGAGACCAATTAAGGATCCCTATTG 2239

QY 1621 TCTCTTTTGGTGTCTATTTTGTCCCTTAAACAAGTCTTTGACAGTGAGAAAAATATTCA 1680
DB 2240 TCTCTTTTGGTGTCTATTTTGTCCCTTAAACAAGTCTTTGACAGTGAGAAAAATATTCA 2299

QY 1681 ATAAACATATCCCTGTGCGCTTATTAATCTTGAACACCTTGCATGAATGAGAGATCC 1740
DB 2300 ATAAACATATCCCTGTGCGCTTATTAATCTTGAACACCTTGCATGAATGAGAGATCC 2359

QY 1741 ACAAGAAAACCTGATGACAACTGTCTTATTTTAACTTATTTGATCATTAAGTTGTAA 1800
DB 2360 ACAAGAAAACCTGATGACAACTGTCTTATTTTAACTTATTTGATCATTAAGTTGTAA 2419

QY 1801 AGAGTAAATAATGTACTCATGTATGATTAATTTTATTAATTTTGGGCTTAAG 1860
DB 2420 AGAGTAAATAATGTACTCATGTATGATTAATTTTATTAATTTTGGGCTTAAG 2479

QY 1861 ATTTTATTAACATGATTCCTTTCTGATATATGAAATGAGCTCTCAAGCTTCATA 1920
 DB 2480 ATTTTATTAACATGATTCCTTTCTGATATATGAAATGAGCTCTCAAGCTTCATA 2539
 QY 1921 AATTATTAACATGATTCCTTTCTGATATATGAAATGAGCTCTCAAGCTTCATA 1980
 DB 2540 AATTATTAACATGATTCCTTTCTGATATATGAAATGAGCTCTCAAGCTTCATA 2599
 QY 1981 GGTCTCTCAAGATTCCTTTCTGATATATGAAATGAGCTCTCAAGCTTCATA 2040
 DB 2600 GGTCTCTCAAGATTCCTTTCTGATATATGAAATGAGCTCTCAAGCTTCATA 2659
 QY 2041 TGGCTCTCAAGATTCCTTTCTGATATATGAAATGAGCTCTCAAGCTTCATA 2100
 DB 2660 TGGCTCTCAAGATTCCTTTCTGATATATGAAATGAGCTCTCAAGCTTCATA 2719
 QY 2101 TACCAATTAACATGATTCCTTTCTGATATATGAAATGAGCTCTCAAGCTTCATA 2160
 DB 2720 TACCAATTAACATGATTCCTTTCTGATATATGAAATGAGCTCTCAAGCTTCATA 2779
 QY 2161 GGAATTAACATG 2172
 DB 2780 GGAATTAACATG 2791

RESULT 4
 AAC55312
 ID AAC55312 standard; cDNA; 2818 BP.
 AC AAC55312;
 XX
 DT 05-FEB-2001 (first entry)
 XX
 DE Human activation-induced cytidine deaminase encoding cDNA SEQ ID NO:7.
 XX
 KM Activation-induced cytidine deaminase; AID; cytidine deaminase;
 KM immune related disease; allergy; allergic disease; antiallergic;
 KM antianemic; antiasclastic; ophthalmological; anti-HIV; dermatological;
 KM gene therapy; B cell associated immune system disorder; food allergy;
 KM immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
 KM IGA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;
 KM drug allergy; allergic rhinitis; Rosen disease; Digorge disease; AIDS;
 KM ataxia telangiectasia; common variable immunodeficiency disorder;
 KM major histocompatibility class II deficiency disease;
 KM auto immunodeficiency syndrome; IgG subclass selection disorder; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 80..676
 FT /tag= a
 FT /product= "activation-induced cytidine deaminase"
 XX
 PN WO200058480-A1.
 XX
 PD 05-OCT-2000.
 XX
 PF 28-MAR-2000; 2000WO-JP001918.
 XX
 PR 29-MAR-1999; 99JP-00087192.
 PR 24-JUN-1999; 99JP-00178999.
 PR 27-DEC-1999; 99JP-00371382.
 XX
 PA (NIB) JAPAN TOBACCO INC.
 PA (HONO/) HONO T.
 PI Honjo T, Muramatsu M;
 XX
 DR WPI, 2000-611715/58.
 DR F-PSDB; AAB24198.
 XX
 PT Nucleic acid encoding activation induced cytidine deaminase, useful as a
 target for drug development for immune-related diseases including

PT allergies.
 XX
 PS Claim 3; Page 135-139; 174p; Japanese.
 XX
 CC The present sequence encodes human activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic,
 CC antiasclastic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorders, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
 CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, Digorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, MHC (major
 CC histocompatibility class) class II deficiency disease, AIDS (auto
 CC immunodeficiency syndrome), elevated IgE disorder, and IgG subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders

XX
 SQ Sequence 2818 BP; 868 A; 548 C; 626 G; 776 T; 0 U; 0 Other;
 Query Match 100.0%; Score 2172; DB 3; Length 2818;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCTGATAGAGTGTGATGATCTTACGAGACGATTTCTGATCTTGGACCTTGTATGCA 60
 DB 623 CCCCTGATAGAGTGTGATGATCTTACGAGACGATTTCTGATCTTGGACCTTGTATGCA 682
 QY 61 CTTCAGAGAAATGTCACACAGATGAAATGCTGCTGAGACAGTGTATAAACAGT 120
 DB 683 CTTCAGAGAAATGTCACACAGATGAAATGCTGCTGAGACAGTGTATAAACAGT 742
 QY 121 CCTCAGAGCTCTCTGTTTATTTCTCAACTCTCACTTTCTTACGATTTACAGAAA 180
 DB 743 CCTCAGAGCTCTCTGTTTATTTCTCAACTCTCACTTTCTTACGATTTACAGAAA 802
 QY 181 AATATTAATACGATCTTTTAAAGATCTATGCTTAAATATGAGAGACACAGT 240
 DB 803 AATATTAATACGATCTTTTAAAGATCTATGCTTAAATATGAGAGACACAGT 862
 QY 241 CTGCGCAGGAGAGTGTGATGATTTGCTGACGTTTAAAGCAATGCCCCACTGTGGA 300
 DB 863 CTGCGCAGGAGAGTGTGATGATTTGCTGACGTTTAAAGCAATGCCCCACTGTGGA 922
 QY 301 ATACAGAACTGAGAGAGCTGGAGAGATCTTAAAGTCAAGCTTTCTATGACTTTTA 360
 DB 923 ATACAGAACTGAGAGAGCTGGAGAGATCTTAAAGTCAAGCTTTCTATGACTTTTA 982
 QY 361 GGTAGAGTGAAGCAGAGAGTGAATCTTAAAGAGATGGTGAAGAGATCAAAATGTTT 420
 DB 983 GGTAGAGTGAAGCAGAGAGTGAATCTTAAAGAGATGGTGAAGAGATCAAAATGTTT 1042
 QY 421 TATCAACATCTTATTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 480
 DB 1043 TATCAACATCTTATTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1102
 QY 481 TCTATCTTTTCCCTGAGCTTAACTTCAAGTAAACAAACCTTCCATCAGGCGCATGA 540
 DB 1103 TCTATCTTTTCCCTGAGCTTAACTTCAAGTAAACAAACCTTCCATCAGGCGCATGA 1162
 QY 541 TCTATAGAGCTCTTATAGAGATCTGGGATGATTTGACCCCAACCATCTTCCAAA 600
 DB 1163 TCTATAGAGCTCTTATAGAGATCTGGGATGATTTGACCCCAACCATCTTCCAAA 1222
 QY 601 GCATTAATATCCATATGAGCTGTATGTTTATATCAGAGAGAGATGTTTATGTTTG 660
 DB 1223 GCATTAATATCCATATGAGCTGTATGTTTATATCAGAGAGAGATGTTTATGTTTG 1282
 QY 661 TACAAAAGAGATTTGTTATGGGTGGGAGTGAAGGTATACCATGATGTCACCTTCA 720

| | | | |
|----|------|---|------|
| Dp | 1283 | TACAAAAGAAATTGTTATGGGTGGGACATGAGGTTATAGACATGCAATGGTCACTTCAA | 1342 |
| Qy | 721 | GCTACTTTAAATAAGAGATCTTAAATGGGACGAGAGACTGTGAAACAAGACCCCTATATA | 780 |
| Dp | 1343 | GCATCTTTAAATAAAGGATCTTAAATGGGACGAGAGACTGTGAAACAAGACCCCTATATA | 1402 |
| Qy | 781 | TGGGTGATGTCTGAAGTGAATACTCTTCTGGAACGCAACTCTTTTAAAGAAATCCCT | 840 |
| Dp | 1403 | TGGGTGATGTCTGAAGTGAATACTCTTCTGGAACGCAACTCTTTTAAAGAAATCCCT | 1462 |
| Qy | 841 | AATTTAGAAACCCCAACCACTTCACATATCATATAATAGCAACAATTGGAAAGAACTTG | 900 |
| Dp | 1463 | AATTTAGAAACCCCAACCACTTCACATATCATATAATAGCAACAATTGGAAAGAACTTG | 1522 |
| Qy | 901 | CTTGAATGTGGGGAGAGGAAATCTATGGCTCTCGGGGTCTCTTCAATCTCGAATATG | 960 |
| Dp | 1523 | CTTGAATGTGGGGAGAGGAAATCTATGGCTCTCGGGGTCTCTTCAATCTCGAATATG | 1582 |
| Qy | 961 | CCAATCAAGSTCAAGSTTTGTCTACATTTTGTATGTGTGATGTCTTCCCAAGSTTAT | 1020 |
| Dp | 1583 | CCAATCAAGSTCAAGSTTTGTCTACATTTTGTATGTGTGATGTCTTCCCAAGSTTAT | 1642 |
| Qy | 1021 | TAACTATATAAGAGATGTGTACAAAACAGATGATAAAGCTGGCAACCGTGACACGC | 1080 |
| Dp | 1643 | TAACTATATAAGAGATGTGTACAAAACAGATGATAAAGCTGGCAACCGTGACACGC | 1702 |
| Qy | 1081 | TCATATGTTCTAGCTGCTTGGAGGTTGAGAGGAGATGGCTGMAACAAGGTGTTCAA | 1140 |
| Dp | 1703 | TCATATGTTCTAGCTGCTTGGAGGTTGAGAGGAGATGGCTGMAACAAGGTGTTCAA | 1762 |
| Qy | 1141 | GGCCAGCCTGGGCAACATPACAAGTCCGTCTCAAAAAAATAAAAAAATAAGAA | 1200 |
| Dp | 1763 | GGCCAGCCTGGGCAACATPACAAGTCCGTCTCAAAAAAATAAAAAAATAAGAA | 1822 |
| Qy | 1201 | GAGAGAGGGCCGGGCGTGTGGCTCAACGCTGTATCCCAAGCACTTTGGAGGCGGAGCC | 1260 |
| Dp | 1823 | GAGAGAGGGCCGGGCGTGTGGCTCAACGCTGTATCCCAAGCACTTTGGAGGCGGAGCC | 1882 |
| Qy | 1261 | GGGCGGATCACTGTGTGTCAGAGATTGAGACACAGCTCGGCCAACAATGGCAAAACCCGT | 1320 |
| Dp | 1883 | GGGCGGATCACTGTGTGTCAGAGATTGAGACACAGCTCGGCCAACAATGGCAAAACCCGT | 1942 |
| Qy | 1321 | CTGTATCTCAAAATGCAAAATTTAGTCAGAGCGTGTGATGAGGCACTGTATATCCAGCTAC | 1380 |
| Dp | 1943 | CTGTATCTCAAAATGCAAAATTTAGTCAGAGCGTGTGATGAGGCACTGTATATCCAGCTAC | 2002 |
| Qy | 1381 | TTGGAGGCTGAGGACGAGAATGCTTGAACCCAGAGGTGAGAGTTGCACTAAGCTGA | 1440 |
| Dp | 2003 | TTGGAGGCTGAGGACGAGAATGCTTGAACCCAGAGGTGAGAGTTGCACTAAGCTGA | 2062 |
| Qy | 1441 | GATCGTCCGCTGCACTTCAGCTGTGGCGACPAAGCAAGCACTCTGTCTCGAATAAAAAA | 1500 |
| Dp | 2063 | GATCGTCCGCTGCACTTCAGCTGTGGCGACPAAGCAAGCACTCTGTCTCGAATAAAAAA | 2122 |
| Qy | 1501 | AAAAAAGAT | 1560 |
| Dp | 2123 | AAAAAAGAT | 2182 |
| Qy | 1561 | TGCAAGAGAAATTTGCTTTATCCAAACAATGTAAAGAGCCAAATPAAGGATCCCTATTG | 1620 |
| Dp | 2183 | TGCAAGAGAAATTTGCTTTATCCAAACAATGTAAAGAGCCAAATPAAGGATCCCTATTG | 2242 |
| Qy | 1621 | TCTCTTTTGTGTCTAATTTGTCCCTAACACTGTCTTTGACAGTGAAGAAAAATTCGA | 1680 |
| Dp | 2243 | TCTCTTTTGTGTCTAATTTGTCCCTAACACTGTCTTTGACAGTGAAGAAAAATTCGA | 2302 |
| Qy | 1681 | ATAACCAATATCCCTGTGCGCTTATTACTAGCAACCTTTCGCAATGAAATAGACAGATCC | 1740 |
| Dp | 2303 | ATAACCAATATCCCTGTGCGCTTATTACTAGCAACCTTTCGCAATGAAATAGACAGATCC | 2362 |
| Qy | 1741 | ACAGGAAACTTGATGACAACTGTCTTAATTTTAACTGTATGTGACATAAGTTGTATA | 1800 |
| Dp | 2363 | ACAGGAAACTTGATGACAACTGTCTTAATTTTAACTGTATGTGACATAAGTTGTATA | 2422 |

| | | | |
|----|------|--|------|
| Qy | 1801 | AGAGTAAAAATGTCCTTCATGATTCATTTATTTATATTTATTTGCGCTAATG | 1860 |
| Db | 2423 | AGAGTAAAAATGTACTCAGATATCATTTATATTTATATTTTGGCGCTAATG | 2482 |
| Qy | 1861 | ATTTTATTTACATGATTTCCCTTTCTGATATATGAAATGAGCTCAAGCTTCATA | 1920 |
| Db | 2483 | ATTTTATTTACATGATTTCCCTTTCTGATATATTTAAATGGAGCTCAAGCTTCATA | 2542 |
| Qy | 1921 | AATTTATACCTTATGGAATATGATTCETATPAACAAGTATGATATGTAACATGGCAGTAT | 1980 |
| Db | 2543 | AATTTATACCTTATGGAATATGATTCETATPAACAAGTATGATATGTAACATGGCAGTAT | 2602 |
| Qy | 1981 | GGTCTACGGAAGCATTTCTCTGATTTTATAGTAACTTTATGACAGCAATTTGCTTC | 2040 |
| Db | 2603 | GGTCTACGGAAGCATTTCTCTGATTTTATAGTAACTTTATGACAGCAATTTGCTTC | 2662 |
| Qy | 2041 | TGCTCTACCTTCATCATCGTTAAATATGATTAATAATTTTGGAACTGTGAGATTAATA | 2100 |
| Db | 2663 | TGCTCTACCTTCATCATCGTTAAATATGATTAATAATTTTGGAACTGTGAGATTAATA | 2722 |
| Qy | 2101 | TACCAATTAATAATATATAAAAGTATTTATGAACTTAATAAATAAATAATCAGTATGAT | 2160 |
| Db | 2723 | TACCAATTAATAATATATAAAAGTATTTATGAACTTAATAAATAAATAATCAGTATGAT | 2782 |
| Qy | 2161 | GGAAATAACTTG 2172 | |
| Db | 2783 | GGAAATAACTTG 2794 | |

```

XX RESULT 5
AC AACS5314
ID AACS5314 standard; DNA; 6564 BP.
XX
XX AACS5314;
XX
XX 05-FEB-2001 (first entry)
DE Human activation-induced cytidine deaminase genomic DNA seq ID No:10.
XX
XX Activation-induced cytidine deaminase; AID; cytidine deaminase;
KM Immune related disease; allergy; allergic disease; antiallergic;
KM antihaemetic; antiasthmatic; ophthalmological; anti-HIV; dermatological;
KM gene therapy; B cell associated immune system disorder; food allergy;
KM immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
KM IgN nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis;
KM drug allergy; allergic rhinitis; Rosen disease; DiGeorge disease; AIDS;
KM ataxia telangiectasia; common variable immunodeficiency disorder;
KM major histocompatibility class II deficiency disease;
KM auto immunodeficiency syndrome; IgG subclass selection disorder; ds.
XX
OS Homo sapiens.
XX
PR WO200058480-A1.
XX
PD 05-OCT-2000.
XX
PF 28-MAR-2000; 2000WO-JP001918.
XX
PR 29-MAR-1999; 99JP-00087192.
PR 24-JUN-1999; 99JP-00178999.
PR 27-DEC-1999; 99JP-00371382.
XX
PA (NISB ) JAPAN TOBACCO INC.
PA (HONJ/) HONJO T.
XX
PI Honjo T, Muramatsu M;
XX
XX WPI; 2000-611715/58.
XX
PR Nucleic acid encoding activation induced cytidine deaminase, useful as a
PT target for drug development for immune-related diseases including
PT allergies.

```


QY 1801 AGAGTAAAAATGTTACTCATGATTCATTTATATATATATTTTGGCTCATG 1860
DB 5541 AGAGTAAAAATGTTACTCATGATTCATTTATATATATATTTTGGCTCATG 5600
QY 1861 ATTTTATATTAACATGATTTCTTTCTGATATATGAAATGAGTCTCAAGCTTCATA 1920
DB 5601 ATTTTATATTAACATGATTTCTTTCTGATATATGAAATGAGTCTCAAGCTTCATA 5660
QY 1921 AATTATTAACCTTTGAAATGATTTCTATATACAGATATGATTTGTAACATGAGTAT 1980
DB 5661 AATTATTAACCTTTGAAATGATTTCTATATACAGATATGATTTGTAACATGAGTAT 5720
QY 1981 GGTGTACAGAGCACTTTCTGTGATTTTGTAGTAACTTTATAGACGAAATTTGCTTC 2040
DB 5721 GGTGTACAGAGCACTTTCTGTGATTTTGTAGTAACTTTATAGACGAAATTTGCTTC 5780
QY 2041 TGGCTCACTTCAATCAGTAAATGATTAATATTTTGAAGCTGTGAAGATTAATA 2100
DB 5781 TGGCTCACTTCAATCAGTAAATGATTAATATTTTGAAGCTGTGAAGATTAATA 5840
QY 2101 TACCAATTAATATATATTAAGATTTATATGATTAATTAATTAATTAATGATATGAT 2160
DB 5841 TACCAATTAATATATATTAAGATTTATATGATTAATTAATTAATTAATGATATGAT 5900
QY 2161 GGAATTAACCTTG 2172
DB 5901 GGAATTAACCTTG 5912

RESULT 6
AAC55339
ID AAC55339 standard; DNA; 11204 BP.
AC AAC55339;
XX
DT 05-FEB-2001 (first entry)
XX
DE Human activation-induced cytidine deaminase genomic DNA SEQ ID NO:35.
XX
DE Human activation-induced cytidine deaminase genomic DNA SEQ ID NO:35.
XX
KW Activation-induced cytidine deaminase; AID; cytidine deaminase;
KW immune related disease; allergy; allergic disease; antiallergic;
KW antianemic; antiaesthetic; ophthalmological; anti-HIV; dermatological;
KW gene therapy; B cell associated immune system disorder; food allergy;
KW immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
KW IGA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;
KW drug allergy; allergic rhinitis; Rosen disease; Digorge disease; AIDS;
KW ataxia telangiectasia; common variable immunodeficiency disorder;
KW major histocompatibility class II deficiency disease;
KW auto immunodeficiency syndrome; Igg subclass selection disorder; ds.
XX
OS Homo sapiens.
XX
PN MO200058480-A1.
XX
XX
PD 05-OCT-2000.
XX
PF 28-MAR-2000; 2000MO-JP001918.
XX
PR 29-MAR-1999; 99JP-00087152.
PR 24-JUN-1999; 99JP-00178999.
PR 27-DEC-1999; 99JP-00371382.
XX
PA (NISB) JAPAN TOBACCO INC.
PA (HONJ) HONJO T.
XX
PI Honjo T, Muramatsu M;
XX
DR WPI; 2000-611715/58.
XX
PT Nucleic acid encoding activation induced cytidine deaminase, useful as a
PT target for drug development for immune-related diseases including
PT allergies.

XX
XX Claim 17; Page 163-170; 174pp; Japanese.
CC The present invention describes an activation-induced cytidine deaminase
CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
CC cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic,
CC antiaesthetic, ophthalmological, anti-HIV and dermatological activities,
CC and can be used in gene therapy. AID polynucleotides are useful in
CC methods for identifying drugs for the treatment of B cell associated
CC immune system disorders, immunodeficiency diseases and allergies, such as
CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
CC globulinemia, allergic dermatitis, allergic colitis, asthma, food allergy,
CC drug allergy, allergic rhinitis, Rosen disease, Digorge disease, ataxia
CC telangiectasia, common variable immunodeficiency disorder, MHC (major
CC histocompatibility class II deficiency disease, AIDS (auto
CC immunodeficiency syndrome), elevated Ige disorder, and Igg subclass
CC selection disorder. The DNA sequences encoding AID may be used for gene
CC therapy and the antibodies to the AID protein may be used for diagnosis
CC and treatment of these disorders. The present sequence represents a
CC genomic DNA sequence of human AID
XX
SQ Sequence 11204 BP; 3305 A; 2273 C; 3253 T; 0 U; 0 Other:
Query Match 100.0%; Score 2172; DB 3; Length 11204;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCCCTGTATGAGGTTGATGATCTTACGAGACGATTTCTGATTTGGACTTTGATAGCAA 60
DB 8956 CCCCTGTATGAGGTTGATGATCTTACGAGACGATTTCTGATTTGGACTTTGATAGCAA 9015
QY 61 CTTCCAGGATGTGACACACGATGAAATATCTCTGTAAGACATGATTAATAAACAAGT 120
DB 9016 CTTCCAGGATGTGACACACGATGAAATATCTCTGTAAGACATGATTAATAAACAAGT 9075
QY 121 CCTTCAAGCTCTCTGTTTATTTCTTCAACTCTCACTTCTTGTAGTTACAGAAAAA 180
DB 9076 CCTTCAAGCTCTCTGTTTATTTCTTCAACTCTCACTTCTTGTAGTTACAGAAAAA 9135
QY 181 AATATTATATACGACTCTTAAATAAAGATCTATGCTTGAATAAGAAAGAACAGGT 240
DB 9136 AATATTATATACGACTCTTAAATAAAGATCTATGCTTGAATAAGAAAGAACAGGT 9195
QY 241 CTGGCCAGGAGAGTGTGTAATTTGTCAGTTTGAATGCAACATTTGCCCTACTGGGA 300
DB 9196 CTGGCCAGGAGAGTGTGTAATTTGTCAGTTTGAATGCAACATTTGCCCTACTGGGA 9255
QY 301 ATTAACGAGACTGACGAGACCTGGGAGCATCTTAAGTTCACAGCTTTCTATGACTTTTA 360
DB 9256 ATTAACGAGACTGACGAGACCTGGGAGCATCTTAAGTTCACAGCTTTCTATGACTTTTA 9315
QY 361 GGTAGATGAGAGCAGAAAGGTAGATCTTAAAGCATGGTGAAGAGATCAATATGTTTAA 420
DB 9316 GGTAGATGAGAGCAGAAAGGTAGATCTTAAAGCATGGTGAAGAGATCAATATGTTTAA 9375
QY 421 TATCAACATCTTTATATTTGATTCATTTGATTAACAGTGTGTAGTAAATTTT 480
DB 9376 TATCAACATCTTTATATTTGATTCATTTGATTAACAGTGTGTAGTAAATTTT 9435
QY 481 TCTATTTCTTTTCCCTTGAAGTTTACTTTCAAGTAAACAAATCTTCCATCAGGCAATCA 540
DB 9436 TCTATTTCTTTTCCCTTGAAGTTTACTTTCAAGTAAACAAATCTTCCATCAGGCAATCA 9495
QY 541 TCTATGAGACTCTGATATGAGATCTGGGATGTTGACCCCAACCATCTCTCCAAA 600
DB 9496 TCTATGAGACTCTGATATGAGATCTGGGATGTTGACCCCAACCATCTCTCCAAA 9555
QY 601 GATTAATATCCATCATGCGCTGTATGTTTATCAAGAGAAAGATGTTTATGTTTG 660
DB 9556 GATTAATATCCATCATGCGCTGTATGTTTATCAAGAGAAAGATGTTTATGTTTG 9615
QY 661 TACAAAAGAAATGTTATGGGTGGGATGAGAGTATAGACCATGATGTCACCTTCAA 720

CC associated with chromosomal aberrations yielding oncogenic fusion
CC proteins (I), treating cancerous cells containing (II) in a heterogenous
CC cell population, treating proliferative diseases associated with mutant
CC protein or cellular protein isoforms (II) dependent on heat shock protein
CC (HSP)-90, or selectively treating cells expressing (II) involving
CC administering HSP90-inhibitor. The method is useful for treating
CC genetically-defined disease with chromosomal aberration yielding
CC oncogenic fusion protein, treating cancerous cells containing fusion
CC protein in heterogenous cell population, treating proliferative disease
CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
CC p53), or selectively treating cells expressing mutant protein or cellular
CC protein isoform in a patient heterozygous for (II). The method is useful
CC for treating a disease e.g. haematopoietic disorder such as T or B cell
CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CML,
CC or a disease characterized by a solid tumour such as papillary thyroid
CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
CC synovial sarcoma. The method is also useful for treating viral
CC infections. This represents the DNA sequence of a chromosome aberration
XX

Sequence 11204 BP; 3305 A; 2273 C; 2373 G; 3253 T; 0 U; 0 Other;

Query Match 100.0%; Score 2172; DB 6; Length 11204;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCTGTATGAGTTGATGACTTACGAGACGCAATTCGTACTTTGGACCTTTGATACGA 60
DB CCCCTGTATGAGTTGATGACTTACGAGACGCAATTCGTACTTTGGACCTTTGATACGA 60
QY 8956
DB 61 CTTCCAGAAATGTCACACAGATGAATATCTGCTGAGACAGTGGATATAAAAGT 120
DB CTTCCAGAAATGTCACACAGATGAATATCTGCTGAGACAGTGGATATAAAAGT 9075
QY 121 CCTCAAGTCTCTCTGTTTATCTTCAACTCTCACTTCTTAGAGTTACAGAAAA 180
DB CCTCAAGTCTCTCTGTTTATCTTCAACTCTCACTTCTTAGAGTTACAGAAAA 180
QY 9076 CTTTCAAGTCTCTCTGTTTATCTTCAACTCTCACTTCTTAGAGTTACAGAAAA 9135
DB CTTTCAAGTCTCTCTGTTTATCTTCAACTCTCACTTCTTAGAGTTACAGAAAA 9135
QY 181 ATATTATATAGACTCTTTTAAAAAGATCTATGCTTGAATAAGAGAGACACAGT 240
DB ATATTATATAGACTCTTTTAAAAAGATCTATGCTTGAATAAGAGAGACACAGT 240
QY 9136 ATATTATATAGACTCTTTTAAAAAGATCTATGCTTGAATAAGAGAGACACAGT 9195
DB ATATTATATAGACTCTTTTAAAAAGATCTATGCTTGAATAAGAGAGACACAGT 9195
QY 241 CTGGCAGGAGAGCTCTGCAATGTGTGAGATTGGAATGCAACATGTCCCTACTGGGA 300
DB CTGGCAGGAGAGCTCTGCAATGTGTGAGATTGGAATGCAACATGTCCCTACTGGGA 300
QY 9196 CTGGCAGGAGAGCTCTGCAATGTGTGAGATTGGAATGCAACATGTCCCTACTGGGA 9255
DB CTGGCAGGAGAGCTCTGCAATGTGTGAGATTGGAATGCAACATGTCCCTACTGGGA 9255
QY 301 ATTAAGAACTGCGAGGACCTGGAGACATCTAAAGTGCACAGTTTTCTATGACTTTA 360
DB ATTAAGAACTGCGAGGACCTGGAGACATCTAAAGTGCACAGTTTTCTATGACTTTA 360
QY 9256 ATTAAGAACTGCGAGGACCTGGAGACATCTAAAGTGCACAGTTTTCTATGACTTTA 9315
DB ATTAAGAACTGCGAGGACCTGGAGACATCTAAAGTGCACAGTTTTCTATGACTTTA 9315
QY 361 GGTAGATGAGAGCAGAGATAGATCTTAAGACATGAGAGATGCAAAATGTTTTTA 420
DB GGTAGATGAGAGCAGAGATAGATCTTAAGACATGAGAGATGCAAAATGTTTTTA 420
QY 9316 GGTAGATGAGAGCAGAGATAGATCTTAAGACATGAGAGATGCAAAATGTTTTTA 9375
DB GGTAGATGAGAGCAGAGATAGATCTTAAGACATGAGAGATGCAAAATGTTTTTA 9375
QY 421 TATCAACATCCTTTATTTATTTGATTCATTTGATTTAACAAGTGTGTTAGATGATTT 480
DB TATCAACATCCTTTATTTATTTGATTCATTTGATTTAACAAGTGTGTTAGATGATTT 480
QY 9376 TATCAACATCCTTTATTTATTTGATTCATTTGATTTAACAAGTGTGTTAGATGATTT 9435
DB TATCAACATCCTTTATTTATTTGATTCATTTGATTTAACAAGTGTGTTAGATGATTT 9435
QY 481 TCTATCTTTTCCCTTGACGTTTACTTCAAGTAAACAAACCTTTCAACAGCCATGA 540
DB TCTATCTTTTCCCTTGACGTTTACTTCAAGTAAACAAACCTTTCAACAGCCATGA 540
QY 9436 TCTATCTTTTCCCTTGACGTTTACTTCAAGTAAACAAACCTTTCAACAGCCATGA 9495
DB TCTATCTTTTCCCTTGACGTTTACTTCAAGTAAACAAACCTTTCAACAGCCATGA 9495
QY 541 TCTATAGACCTCCTTAATGAGAGATCTGGGTGATGTGACCCCAACCATCTCCAAA 600
DB TCTATAGACCTCCTTAATGAGAGATCTGGGTGATGTGACCCCAACCATCTCCAAA 600
QY 9496 TCTATAGACCTCCTTAATGAGAGATCTGGGTGATGTGACCCCAACCATCTCCAAA 9555
DB TCTATAGACCTCCTTAATGAGAGATCTGGGTGATGTGACCCCAACCATCTCCAAA 9555
QY 601 GCATTATATCCATCATGCGCTGATGTTTAAATCAGCAGACAGACATGTTTATGTTG 660
DB GCATTATATCCATCATGCGCTGATGTTTAAATCAGCAGACAGACATGTTTATGTTG 660
QY 9556 GCATTATATCCATCATGCGCTGATGTTTAAATCAGCAGACAGACATGTTTATGTTG 9615
DB GCATTATATCCATCATGCGCTGATGTTTAAATCAGCAGACAGACATGTTTATGTTG 9615
QY 661 TACAAAGAGAAATGTTATGAGTGGAGATGAGAGATGACATGATGTCACCTTCAA 720
DB TACAAAGAGAAATGTTATGAGTGGAGATGAGAGATGACATGATGTCACCTTCAA 720
QY 9616 TACAAAGAGAAATGTTATGAGTGGAGATGAGAGATGACATGATGTCACCTTCAA 9675
DB TACAAAGAGAAATGTTATGAGTGGAGATGAGAGATGACATGATGTCACCTTCAA 9675

QY 721 GCTACTTTATATAAGATCTTAAATGGGACAGAGAGCTGTGACAAACACCTTAATTA 780
DB GCTACTTTATATAAGATCTTAAATGGGACAGAGAGCTGTGACAAACACCTTAATTA 780
QY 9676 GCTACTTTATATAAGATCTTAAATGGGACAGAGAGCTGTGACAAACACCTTAATTA 9735
DB GCTACTTTATATAAGATCTTAAATGGGACAGAGAGCTGTGACAAACACCTTAATTA 9735
QY 781 TGGGTGATGTCTGAATGACAAATCTCTGGAACGCAACTCTTTAAGAGTCCCT 840
DB TGGGTGATGTCTGAATGACAAATCTCTGGAACGCAACTCTTTAAGAGTCCCT 840
QY 9736 TGGGTGATGTCTGAATGACAAATCTCTGGAACGCAACTCTTTAAGAGTCCCT 9795
DB TGGGTGATGTCTGAATGACAAATCTCTGGAACGCAACTCTTTAAGAGTCCCT 9795
QY 841 AATTATGAAACACCCCAAACTTTCATATCAATTTAGCAAAAGATTTGAAAGAGTGG 900
DB AATTATGAAACACCCCAAACTTTCATATCAATTTAGCAAAAGATTTGAAAGAGTGG 900
QY 9796 AATTATGAAACACCCCAAACTTTCATATCAATTTAGCAAAAGATTTGAAAGAGTGG 9855
DB AATTATGAAACACCCCAAACTTTCATATCAATTTAGCAAAAGATTTGAAAGAGTGG 9855
QY 901 CTTGAATGTTGGGAGAGAGAAATATGATGCTCGTGGGTCTCTTCAATCAGAAATG 960
DB CTTGAATGTTGGGAGAGAGAAATATGATGCTCGTGGGTCTCTTCAATCAGAAATG 960
QY 9856 CTTGAATGTTGGGAGAGAGAAATATGATGCTCGTGGGTCTCTTCAATCAGAAATG 9915
DB CTTGAATGTTGGGAGAGAGAAATATGATGCTCGTGGGTCTCTTCAATCAGAAATG 9915
QY 961 CCAATCAGGTCAAGTTTGTCTACATTTGTATGTGTGATGCTTCCCAAGGTATAT 1020
DB CCAATCAGGTCAAGTTTGTCTACATTTGTATGTGTGATGCTTCCCAAGGTATAT 1020
QY 9916 CCAATCAGGTCAAGTTTGTCTACATTTGTATGTGTGATGCTTCCCAAGGTATAT 9975
DB CCAATCAGGTCAAGTTTGTCTACATTTGTATGTGTGATGCTTCCCAAGGTATAT 9975
QY 1021 TAACTATATAGAGAGTTGTGACAAACAGATGATTAAGCTGGACACCGTGGCACAGC 1080
DB TAACTATATAGAGAGTTGTGACAAACAGATGATTAAGCTGGACACCGTGGCACAGC 1080
QY 9976 TAACTATATAGAGAGTTGTGACAAACAGATGATTAAGCTGGACACCGTGGCACAGC 10035
DB TAACTATATAGAGAGTTGTGACAAACAGATGATTAAGCTGGACACCGTGGCACAGC 10035
QY 1081 TCAATGTTCTACTGCTTGGAGGTTGAGAGAGAGATGCTTGAACACAGTGTGAA 1140
DB TCAATGTTCTACTGCTTGGAGGTTGAGAGAGAGATGCTTGAACACAGTGTGAA 1140
QY 10036 TCAATGTTCTACTGCTTGGAGGTTGAGAGAGAGATGCTTGAACACAGTGTGAA 10095
DB TCAATGTTCTACTGCTTGGAGGTTGAGAGAGAGATGCTTGAACACAGTGTGAA 10095
QY 1141 GGCACAGCTGGGCAACATTAACAAGATCTCTCTCAAAAAAAGAGAGAGAGAGAGAG 1200
DB GGCACAGCTGGGCAACATTAACAAGATCTCTCTCAAAAAAAGAGAGAGAGAGAGAG 1200
QY 10096 GGCACAGCTGGGCAACATTAACAAGATCTCTCTCAAAAAAAGAGAGAGAGAGAGAG 10155
DB GGCACAGCTGGGCAACATTAACAAGATCTCTCTCAAAAAAAGAGAGAGAGAGAGAG 10155
QY 1201 GAGAGAGGCGGGGCGTGTGCTGACGCTGTAAATCCAGACATTTGGAGGCGGAGCC 1260
DB GAGAGAGGCGGGGCGTGTGCTGACGCTGTAAATCCAGACATTTGGAGGCGGAGCC 1260
QY 10156 GAGAGAGGCGGGGCGTGTGCTGACGCTGTAAATCCAGACATTTGGAGGCGGAGCC 10215
DB GAGAGAGGCGGGGCGTGTGCTGACGCTGTAAATCCAGACATTTGGAGGCGGAGCC 10215
QY 1261 GGGCGGATCACTGTGTGTGAGAGAGTTGAGACACAGCTGGCCAAATGCGGAGAGAG 1320
DB GGGCGGATCACTGTGTGTGAGAGAGTTGAGAGAGAGTTGAGAGAGAGTTGAGAGAG 1320
QY 10216 GGGCGGATCACTGTGTGTGAGAGAGTTGAGAGAGAGTTGAGAGAGAGTTGAGAGAG 10275
DB GGGCGGATCACTGTGTGTGAGAGAGTTGAGAGAGAGTTGAGAGAGAGTTGAGAGAG 10275
QY 1321 CTGTACTCAAAATGCAAAAAATTAGCAGGCGGTGAGAGAGGACCTGTATCCAGCTAC 1380
DB CTGTACTCAAAATGCAAAAAATTAGCAGGCGGTGAGAGAGGACCTGTATCCAGCTAC 1380
QY 10276 CTGTACTCAAAATGCAAAAAATTAGCAGGCGGTGAGAGAGGACCTGTATCCAGCTAC 10335
DB CTGTACTCAAAATGCAAAAAATTAGCAGGCGGTGAGAGAGGACCTGTATCCAGCTAC 10335
QY 1381 TTGGAGGCTGAGGAGAGAGAAATGCTTGAACCCAGAGGTGAGAGTTGACATAGCTGA 1440
DB TTGGAGGCTGAGGAGAGAGAAATGCTTGAACCCAGAGGTGAGAGTTGACATAGCTGA 1440
QY 10336 TTGGAGGCTGAGGAGAGAGAAATGCTTGAACCCAGAGGTGAGAGTTGACATAGCTGA 10395
DB TTGGAGGCTGAGGAGAGAGAAATGCTTGAACCCAGAGGTGAGAGTTGACATAGCTGA 10395
QY 1441 GATCGTCCGTTGACCTCGAGCTGGGCGACAAAGAGACACTCTGTCTCAAAAAA 1500
DB GATCGTCCGTTGACCTCGAGCTGGGCGACAAAGAGACACTCTGTCTCAAAAAA 1500
QY 10396 GATCGTCCGTTGACCTCGAGCTGGGCGACAAAGAGACACTCTGTCTCAAAAAA 10455
DB GATCGTCCGTTGACCTCGAGCTGGGCGACAAAGAGACACTCTGTCTCAAAAAA 10455
QY 1501 AAAAAAGAT 1560
DB AAAAAAGAT 1560
QY 10456 AAAAAAGAT 10515
DB AAAAAAGAT 10515
QY 1561 TGAAGAGAAATGTGTTTATATCAACAAATGTAAGAGCCAAATAGGAGATCCATTTTG 1620
DB TGAAGAGAAATGTGTTTATATCAACAAATGTAAGAGCCAAATAGGAGATCCATTTTG 1620
QY 10516 TGAAGAGAAATGTGTTTATATCAACAAATGTAAGAGCCAAATAGGAGATCCATTTTG 10575
DB TGAAGAGAAATGTGTTTATATCAACAAATGTAAGAGCCAAATAGGAGATCCATTTTG 10575
QY 1621 TCTCTTTTGTCTATTTTGTCCCTTAACAATGTCTTTGACAGTGAAGAAAAATATTCGA 1680
DB TCTCTTTTGTCTATTTTGTCCCTTAACAATGTCTTTGACAGTGAAGAAAAATATTCGA 1680
QY 10576 TCTCTTTTGTCTATTTTGTCCCTTAACAATGTCTTTGACAGTGAAGAAAAATATTCGA 10635
DB TCTCTTTTGTCTATTTTGTCCCTTAACAATGTCTTTGACAGTGAAGAAAAATATTCGA 10635
QY 1681 ATTAACATATCCCTGTGCGGTTATTAAGTACAGCACTTTGCAATGAAGATGAGCATCC 1740
DB ATTAACATATCCCTGTGCGGTTATTAAGTACAGCACTTTGCAATGAAGATGAGCATCC 1740
QY 10636 ATTAACATATCCCTGTGCGGTTATTAAGTACAGCACTTTGCAATGAAGATGAGCATCC 10695
DB ATTAACATATCCCTGTGCGGTTATTAAGTACAGCACTTTGCAATGAAGATGAGCATCC 10695
QY 1741 ACAGAGAACTGAG 1800
DB ACAGAGAACTGAG 1800
QY 10696 ACAGAGAACTGAG 10755
DB ACAGAGAACTGAG 10755

| | | |
|----|---|------------------|
| XX | 08-NOV-2000; | 2000US-0246526P. |
| PR | 08-NOV-2000; | 2000US-0246527P. |
| PR | 08-NOV-2000; | 2000US-0246528P. |
| PR | 08-NOV-2000; | 2000US-0246529P. |
| PR | 08-NOV-2000; | 2000US-0246532P. |
| PR | 08-NOV-2000; | 2000US-0246609P. |
| PR | 08-NOV-2000; | 2000US-0246610P. |
| PR | 08-NOV-2000; | 2000US-0246611P. |
| PR | 08-NOV-2000; | 2000US-0246613P. |
| PR | 17-NOV-2000; | 2000US-0249207P. |
| PR | 17-NOV-2000; | 2000US-0249208P. |
| PR | 17-NOV-2000; | 2000US-0249209P. |
| PR | 17-NOV-2000; | 2000US-0249210P. |
| PR | 17-NOV-2000; | 2000US-0249211P. |
| PR | 17-NOV-2000; | 2000US-0249212P. |
| PR | 17-NOV-2000; | 2000US-0249213P. |
| PR | 17-NOV-2000; | 2000US-0249214P. |
| PR | 17-NOV-2000; | 2000US-0249215P. |
| PR | 17-NOV-2000; | 2000US-0249216P. |
| PR | 17-NOV-2000; | 2000US-0249217P. |
| PR | 17-NOV-2000; | 2000US-0249218P. |
| PR | 17-NOV-2000; | 2000US-0249244P. |
| PR | 17-NOV-2000; | 2000US-0249245P. |
| PR | 17-NOV-2000; | 2000US-0249264P. |
| PR | 17-NOV-2000; | 2000US-0249265P. |
| PR | 17-NOV-2000; | 2000US-0249297P. |
| PR | 17-NOV-2000; | 2000US-0249299P. |
| PR | 17-NOV-2000; | 2000US-0249300P. |
| PR | 01-DEC-2000; | 2000US-0250130P. |
| PR | 01-DEC-2000; | 2000US-0250319P. |
| PR | 05-DEC-2000; | 2000US-0251030P. |
| PR | 05-DEC-2000; | 2000US-0251988P. |
| PR | 05-DEC-2000; | 2000US-0256719P. |
| PR | 06-DEC-2000; | 2000US-0251479P. |
| PR | 08-DEC-2000; | 2000US-0251856P. |
| PR | 08-DEC-2000; | 2000US-0251868P. |
| PR | 08-DEC-2000; | 2000US-0251869P. |
| PR | 08-DEC-2000; | 2000US-0251899P. |
| PR | 08-DEC-2000; | 2000US-0251907P. |
| PR | 11-DEC-2000; | 2000US-0254097P. |
| PR | 05-JAN-2001; | 2001US-0259678P. |
| PA | (HUMA-) HUMAN GENOME SCI INC. | |
| XX | Rosen CA, Barash SC, Ruben SW; | |
| PT | WPI; 2001-483426/52. | |
| DR | Nucleic acids encoding human immune/hematopoietic antigen polypeptides, | |
| XX | useful for preventing, diagnosing and/or treating cancers and metastasis | |
| XX | Disclosure; SEQ ID NO 35900; 3071pp + Sequence Listing; English. | |
| PS | AAX54951 to AAK64702 encode the human immune/haematopoietic antigen (I) | |
| CC | amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic | |
| CC | activity, and can be used in gene therapy and vaccine production. (II) | |
| CC | proteins and polynucleotides may be used in the prevention, diagnosis and | |
| CC | treatment of diseases associated with inappropriate (I) expression. For | |
| CC | example, they may be used to treat disorders associated with decreased | |
| CC | expression by rectifying mutations or deletions in a patient's genome | |
| CC | that affect the activity of (I) by expressing inactive proteins or to | |
| CC | supplement the patients own production of (I). Additionally, (I) | |
| CC | nucleic acids may be used to produce the secreted (I), by inserting the | |
| CC | nucleic acids into a host cell and culturing the cell to express the | |
| CC | protein. (I) proteins and polynucleotides may be used to prevent, | |
| CC | diagnose and treat immune/haematopoietic-related diseases, especially | |
| CC | cancers and cancer metastases of haematopoietic-derived cells. AAK64703 | |
| CC | to AAK87694 represent human immune/haematopoietic antigen genomic | |
| CC | sequences from the present invention. AAK54942 to AAK54950 and AAM82169 | |
| CC | represent sequences used in the exemplification of the present invention | |
| XQ | Sequence 1665 BP; 497 A; 343 C; 365 G; 460 T; 0 U; 0 Other; | |

Query Match 68.5%; Score 1488; DB 4; Length 1665;

| | |
|----|--|
| | Best local Similarity 100.0%; Pred. No. 66-268; Matches 1468; Conservative 0; Mismatches 0; Indels 0; Gaps 0 |
| QY | 1 CCCCTGTAATGAGTTGATGACTTACGACACGCACTTTCGATCTTTGGACCTTTGATACGA 60 |
| Db | 178 CCCCTGTAATGAGTTGATGACTTACGACACGCACTTTCGATCTTTGGACCTTTGATACGA 237 |
| QY | 61 CTTCCAGGAATGTACACACGATGAAATATCTCGCTGAAAGACAGTGAATAAAAAAGT 120 |
| Db | 238 CTTCCAGGAATGTACACACGATGAAATATCTCGCTGAAAGACAGTGAATAAAAAAGT 297 |
| QY | 121 CCTCAAGTCTCTCTGTTTTTATCTCAACTCTCACTTCTTAGAGTTACAGAAAA 180 |
| Db | 298 CCTCAAGTCTCTCTGTTTTTATCTCAACTCTCACTTCTTAGAGTTACAGAAAA 357 |
| QY | 181 AATATTATATACGACTCTTTAAAAAGTCTATGTCCTTGAAATAGAGAGAACACAGGT 240 |
| Db | 358 AATATTATATACGACTCTTTAAAAAGTCTATGTCCTTGAAATAGAGAGAACACAGGT 417 |
| QY | 241 CTGGCCAGGAGACGTGCTGCATTTGCTGCATTTTGAATGCAACATTGTCCTTACGGA 300 |
| Db | 418 CTGGCCAGGAGACGTGCTGCATTTGCTGCATTTTGAATGCAACATTGTCCTTACGGA 477 |
| QY | 301 ATPAACAGAACTGCAGGACCTGGGAGCATCTTAAAGTCAACGTTTTTCTATGACTTTTA 360 |
| Db | 478 ATPAACAGAACTGCAGGACCTGGGAGCATCTTAAAGTCAACGTTTTTCTATGACTTTTA 537 |
| QY | 361 GGTAGATGAGACGAAAGGTAGATCTTAAAAAGCATGCTGAGAGAGATCAAAATGTTTTTA 420 |
| Db | 538 GGTAGATGAGACGAAAGGTAGATCTTAAAAAGCATGCTGAGAGAGATCAAAATGTTTTTA 597 |
| QY | 421 TATCAACATCCTTTATATTTGATTCATTGAGTTAACAGTGTGTAGTATGATTTT 480 |
| Db | 598 TATCAACATCCTTTATATTTGATTCATTGAGTTAACAGTGTGTAGTATGATTTT 657 |
| QY | 481 TCTATCTTTCCCTGACGTTTACTTCAAGTACACAACTCTTCATCAGGCGCATGA 540 |
| Db | 658 TCTATCTTTCCCTGACGTTTACTTCAAGTACACAACTCTTCATCAGGCGCATGA 717 |
| QY | 541 TCTATGAGACCTCTTAATGAGAGATCTGGGATTTGTGAACCCCAACCATCTCCAA 600 |
| Db | 718 TCTATGAGACCTCTTAATGAGAGATCTGGGATTTGTGAACCCCAACCATCTCCAA 777 |
| QY | 601 GCATTATATATCAATCAATGCGCTGATGTTTATACGACGAAAGCATGTTTTATGTTTG 660 |
| Db | 778 GCATTATATATCAATCAATGCGCTGATGTTTATACGACGAAAGCATGTTTTATGTTTG 837 |
| QY | 661 TACAAAABAATGTTATGCGGTGGGGAATGCAAGTATACCATGATGTCACCTTCAA 720 |
| Db | 838 TACAAAABAATGTTATGCGGTGGGGAATGCAAGTATACCATGATGTCACCTTCAA 897 |
| QY | 721 GCTACTTAATTAAGAAGATCTTAAAAATGGGACAGAGACTGTGACAAAGCACCTTAATA 780 |
| Db | 898 GCTACTTTTAATTAAGAAGATCTTAAAAATGGGACAGAGACTGTGACAAAGCACCTTAATA 957 |
| QY | 781 TGGGTGATGTCGTAGATGACAAATCTTCTGAAAACGCAAACTCTTTTAAGGAAGTCCT 840 |
| Db | 958 TGGGTGATGTCGTAGATGACAAATCTTCTGAAAACGCAAACTCTTTTAAGGAAGTCCT 1011 |
| QY | 841 AATTAGAAAACCCACAAACTTCACATACATATTAAGCAACAAATGGAAGAAAGTTG 900 |
| Db | 1018 AATTAGAAAACCCACAAACTTCACATACATATTAAGCAACAAATGGAAGAAAGTTG 1077 |
| QY | 901 CTTGAATGTTGGGAGAGAAAATCTATTTGGCTCTCGTGGGTCCTTCACTCAGAAATG 960 |
| Db | 1078 CTTGAATGTTGGGAGAGAAAATCTATTTGGCTCTCGTGGGTCCTTCACTCAGAAATG 1133 |
| QY | 961 CCAATCAGGTCAAGGTTTGCTACATTTTGTATGTGTATGCTTCCCAAGGATAT 1020 |
| Db | 1138 CCAATCAGGTCAAGGTTTGCTACATTTTGTATGTGTATGCTTCCCAAGGATAT 1197 |
| QY | 1021 TTACTATTAAGAAGATTGTGACAAACAAATGATTAAGCTGCCAACCCTGGCACAGC 1088 |

Db 1341 GCTACTTATAAAGATCTTAAATGGCAGAGGACTGTGACACAGACCCCTAATTA 1400
QY 781 TGGGTGAGTCTGAGTAGCAAAATCTTGTGAAACGCAACTTTTAAAGAACTCCT 840
Db 1401 TGGGTGAGTCTGAGTAGCAAAATCTTGTGAAACGCAACTTTTAAAGAACTCCT 1460
QY 841 AATTGAAACACCCCAACTTCACATATCATTAATTAGCAACATTGAGAGAGTTG 900
Db 1461 AATTGAAACACCCCAACTTCACATATCATTAATTAGCAACATTGAGAGAGTTG 1520
QY 901 CTTGAATGTTGGGAGAGAGAAA 923
Db 1521 CTTGAATGTTGGGAGAGAGAAA 1543
RESULT 10
AAK81089
ID AAK81089 standard; DNA; 574 BP.
XX AAK81089;
AC AAK81089;
XX
XX 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35901.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN M0200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US001354.
XX
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216474P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218920P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225113P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225577P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232337P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236377P.
PR 29-SEP-2000; 2000US-0236378P.
PR 29-SEP-2000; 2000US-0236386P.
PR 29-SEP-2000; 2000US-0236389P.
PR 29-SEP-2000; 2000US-0236390P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241211P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244647P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246533P.
PR 08-NOV-2000; 2000US-0246534P.
PR 08-NOV-2000; 2000US-0246545P.
PR 08-NOV-2000; 2000US-0246546P.
PR 08-NOV-2000; 2000US-0246572P.
PR 08-NOV-2000; 2000US-0246582P.
PR 08-NOV-2000; 2000US-0246583P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.

Query Match 19.8%; Score 429; DB 4; Length 429;
Best Local Similarity 100.0%; Pred. No. 1e-70;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 764 ACAGACACCCCTAATAGGCTGATGCTGAGTAGCAATCTCTGAGAGCAAACT 823
DB 1 ACAAGACACCCCTAATAGGCTGATGCTGAGTAGCAATCTCTGAGAGCAAACT 60

QY 824 CTTTAAAGAAAGTCCCTAATTAGAAACACCCCAAACTTCAATATCATATTAGCAA 883
DB 61 CTTTAAAGAAAGTCCCTAATTAGAAACACCCCAAACTTCAATATCATATTAGCAA 120

QY 884 CAATTGAGAGAGTCTGATGATGTTGGGAGAGAGAAATCTATTGGCTCTGAGGTC 943
DB 121 CAATTGAGAGAGTCTGATGATGTTGGGAGAGAGAAATCTATTGGCTCTGAGGTC 180

QY 944 TCTTCATCTCGAAATGCGAATCGCTCAAGTTTGTATCATTTGTATGTGTATGC 1003
DB 181 TCTTCATCTCGAAATGCGAATCGCTCAAGTTTGTATGTGTATGTGTATGC 240

QY 1004 TTTCCCAAGAGTATATTAATATATAGAGAGTTGTGACAAAACAGATGATAAGCTG 1063
DB 241 TTTCCCAAGAGTATATTAATATATAGAGAGTTGTGACAAAACAGATGATAAGCTG 300

QY 1064 CGAACCGTGGCAGACGCTCATATGTTCTAGCTGCTTGGAGAGTTGAGAGAGAGATGGCT 1123
DB 301 CGAACCGTGGCAGACGCTCATATGTTCTAGCTGCTTGGAGAGTTGAGAGAGAGATGGCT 360

QY 1124 TGAACACAGTGTTCAGAGCCAGCTGGGCAACATAACAGATCTCTCTCAAAAAA 1183
DB 361 TGAACACAGTGTTCAGAGCCAGCTGGGCAACATAACAGATCTCTCTCAAAAAA 420

QY 1184 AAAAAAAAA 1192
DB 421 AAAAAAAAA 429

RESULT 12
ID AAL02789 standard; DNA; 13862 BP.
XX AAL02789;
XX
XX
XX 21-NOV-2001 (first entry)
XX
XX Human reproductive system related antigen DNA SEQ ID NO: 5477.
XX
XX Human; reproductive system related antigen; reproductive system disorder;
XX
XX cancer; gene therapy; ds.
XX
XX Homo sapiens.
XX
XX
XX WO20015320-A2.
XX
XX
XX 02-AUG-2001.
XX
XX
XX 17-JAN-2001; 2001WO-US001339.
XX
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
XX
XX 30-JUN-2000; 2000US-0215135P.
XX
XX 07-JUL-2000; 2000US-0216647P.
XX
XX 07-JUL-2000; 2000US-0216809P.
XX
XX 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225265P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225785P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227099P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234979P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Barash SC, Ruben SM;

XX DR WPI; 2001-465570/50.

XX PT Isolated nucleic acid molecule encoding a reproductive system antigen is used in preventing, treating or ameliorating a medical condition.

XX PS Disclosure; SEQ ID NO 5477; 1297pp + Sequence Listing; English.

XX CC The present invention provides the protein and coding sequences of a number of human reproductive system related antigens. These can be used in the prevention and treatment of reproductive system disorders, including cancer. The present sequence is a genomic sequence encoding a protein of the invention

XX SQ Sequence 13862 BP; 3188 A; 3560 C; 4465 G; 2649 T; 0 U; 0 Other;

Query Match 13.6%; Score 295.2; DB 4; Length 13862;
 Best Local Similarity 74.6%; Pred. No. 1.5e-45;
 Matches 397; Conservative 0; Mismatches 133; Indels 2; Gaps 2;

QY 1045 AACAGATGATTAAGCTGCAACCGTGGCACAGCCTAGTATTCTAGCTCTGGAGG 1104
 DB 1508 AAACCTTAATTAAGCTGCTGGGCTCTGTGGCTCAGGCCCTGTATCCACGACTTTGGAGG 1567
 QY 1105 TTGAGGAGGAGATGCTTTGAACAACAGGTGTTCAAGGCCAGCTGGCAACATTAACAG 1164
 DB 1568 CTGAGGACGAGGATCATCTTGAGGTCAAGATTGACACAGCTGGCCAACTAGGCCAA 1627
 QY 1165 ATCCGTCTCTCAAAAAAAAAAAAAAAAAAGAAAG- GAGAGGGCCGGCGTGGCC 1223
 DB 1628 ACCCATCTCTAATAAATCAAAAGTTACCCAGGCTGAGCCGGGATGAGTGGGTGGC 1687
 QY 1224 TCAGCCTGTATCCACGACTTTGGAGGCCGAGCCGGGAGATCACCCTGTGTCAGGA 1283
 DB 1688 TATGCTCTGTATCCACGACTTTGGAGGCCGAGTGGGAGTCACTTACGTACAGG 1747
 QY 1284 GTTGAACACAGCCTGGCCAAATGCAAAACCCGCTGTACTCAAAATGCAAAATTA 1343
 DB 1748 GTTCAAGACCAAGCCTGACCAACATAGCAAAACCGTCTTATTAATAAATAAATTA 1807
 QY 1344 GCCAGCGTGTGAGGACCTGTATCCAGCTACTTGGAGGCTGAGGACAGAGAT 1403
 DB 1808 GCCAGGCTTATGTAGGAGCGCTTATATCCAGCTACTGGAAGGCTGAGGACAGAGAT 1867
 QY 1404 CGCTTGACCCAGAGGTGAGGTGCACTAGCTGAGTGAATGTCGCTGACTCCAGCC 1463
 DB 1868 CGTTGAAGCCAGAGGCGGAGTTGACCAAGCCGAGATGTGCACTGCACTCCAGCC 1927
 QY 1464 TGGGCGACAGAGCAAGCTGTCTCAGAAAAAAGAAAAAGAGAGAGAGAA 1523
 DB 1928 TGGGCGATTA-AGCAGACTCTGTCTCAGAAAAAAGAAAAAGAAAAAGAAAAA 1986
 QY 1524 AGAGACATATTTGGAG 1575
 DB 1987 AGAAAAAATGACCGAGCGTGGGATGCACTGATGCTAGTACTTGGG 2038

RESULT 13
 AAL07516
 ID AAL07516 standard; DNA; 13862 BP.
 XX
 AC AAL07516;
 XX
 DT 21-NOV-2001 (first entry)
 XX
 DE Human reproductive system related antigen DNA SEQ ID NO: 10204.
 XX
 KW Human; reproductive system related antigen; reproductive system disorder;
 KW cancer; gene therapy; ds.
 XX
 OS Homo sapiens.
 OS
 PN WO200155320-A2.
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001339.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216800P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.

| | |
|----|--|
| PA | (HUMA-HUMAN GENOME SCI INC. |
| XX | |
| PI | Rosen CA, Barash SC, Ruben SM; |
| XX | |
| DR | WPI, 2001-488786/53. |
| XX | |
| PT | New isolated ovarian and/or breast cancer related nucleic acids and |
| PT | polypeptides, useful for diagnosing, treating and/or preventing human |
| PT | diseases and disorders, particularly ovarian and/or breast cancer. |
| XX | |
| PS | Disclosure; SEQ ID NO 1003; 577bp + Sequence Listing; English. |
| XX | |
| CC | The invention relates to novel genes (ABAO7454-ABAO8224) and proteins |
| CC | (ABAO10743-ABAO10980) useful for preventing, treating or ameliorating |
| CC | medical conditions e.g., by protein or gene therapy. The gene are |
| CC | isolated from a range of human tissues disclosed in the specification. |
| CC | The nucleic acids, proteins, antibodies and (ant)agonists are useful in |
| CC | the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and |
| CC | ovarian cancer and other cancers of the adrenal gland, bone, bone marrow |
| CC | breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune |
| CC | disorders e.g. Addison's disease, allergies, autoimmune haemolytic |
| CC | anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, |

| | | | | |
|-----------------------|--------------|--------------------|-----------------|---------------|
| Query Match | 13.6%; | Score 295.2; | DB 4; | Length 13862; |
| Best Local Similarity | 74.6%; | Pred. No. 1.5e-45; | | |
| Matches 397; | Conservative | 0; | Mismatches 133; | Indels 2; |
| | | | | Gaps 2 |

```

RESULT 15
ACF42745.0/c
WP Sequence split into 4 fragments      Locus ACF42745 Accession ACF42745
WP                                     Fragment Name      Begin      End
WP      ACF42745_0                      1          110000
WP      ACF42745_1                     100001     210000
WP      ACF42745_2                     200001     310000
WP      ACF42745_3                     300001     354391
ID      ACF42745 Standard; DNA; 354391 BP.
XX
XX      ACF42745;
AC
XX
XX      29-SHP-2003 (first entry)
XX
XX
DE      Human ALMS1 genomic DNA sequence.
XX
XX      Human; ALMS1; chromosome 2; 2p13; Alstrom disease; retinal dystrophy;
KW      cardiomyopathy; endocrinopathy; diabetes; Alstrom syndrome; cardiac;
KW      ophthalmological; antidiabetic; hepatotropic; nephrotropic; gene therapy;
XX      gene; ds.
XX
OS      Homo sapiens.

```

[illegible]

| | | | |
|----|-------|---|-------|
| Qy | 1360 | GGGACCCGTATATCCCAAGACTCTTTGGGAGCTAGGGAGAGATGCGCTTGAACCCAGAG | 1419 |
| Db | 21204 | CATGCTCTGATATCCCAAGCTACTTTGAAGGCTAGGTGAGTAAAGT-GCTTGAACCTGGAG | 21146 |
| Qy | 1420 | GTGGAGGTTGCAGTAAGCTGAGATCGTGCCGTTGCATCCAGCTGGGCGACAAGAGCAA | 1479 |
| Db | 21145 | GCGGAGGTTTGCAGTAGGCCAAGTCTTACATTGCATTCGACCTGGGCAACAAGCCA | 21088 |
| Qy | 1480 | GACTCTGTCTCAGAAAAAAGAGAGAGAGA | 1516 |
| Db | 21085 | AATCCATCTCAACAAACAAAAAATTTAACTGAGA | 21049 |

```

Search completed: March 12, 2004, 18:08:15
Job time : 874.74 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 17:39:25 ; Search time 161.607 Seconds
(without alignments)
7458.540 Million cell updates/sec

Title: US-09-966-880A-15

Perfect score: 2172
Sequence: 1 cccctgcatgagtgatgata.....agatgcatgataaacttg 2172

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCUTS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|---------|-------------------------------------|
| 1 | 280 | 12.9 | 15297 4 | US-09-817-180-3 Sequence 3, Appl |
| 2 | 280 | 12.9 | 15297 4 | US-10-003-285-3 Sequence 3, Appl |
| 3 | 274.2 | 12.6 | 11811 3 | US-09-078-284-7 Sequence 7, Appl |
| 4 | 269.8 | 12.4 | 43950 4 | US-09-735-934A-3 Sequence 3, Appl |
| 5 | 269.8 | 12.4 | 43950 4 | US-10-060-332-3 Sequence 3, Appl |
| 6 | 269.2 | 12.4 | 99500 4 | US-09-798-096-10 Sequence 10, Appl |
| 7 | 267.4 | 12.3 | 59065 4 | US-09-813-817-3 Sequence 3, Appl |
| 8 | 267.4 | 12.3 | 59065 4 | US-09-978-187-3 Sequence 3, Appl |
| 9 | 265.6 | 12.2 | 9365 4 | US-09-608-285A-8 Sequence 8, Appl |
| 10 | 265.6 | 12.2 | 9365 4 | US-09-350-836B-8 Sequence 8, Appl |
| 11 | 265.6 | 12.2 | 9365 4 | US-09-370-265-8 Sequence 8, Appl |
| 12 | 265.6 | 12.2 | 9365 4 | US-09-557-800C-8 Sequence 8, Appl |
| 13 | 265.6 | 12.2 | 9365 4 | US-09-370-625A-8 Sequence 8, Appl |
| 14 | 265.6 | 12.2 | 14747 4 | US-09-608-285A-42 Sequence 42, Appl |
| 15 | 265.6 | 12.2 | 14747 4 | US-09-557-800C-42 Sequence 42, Appl |
| 16 | 265.6 | 12.2 | 15977 4 | US-09-608-285A-59 Sequence 59, Appl |
| 17 | 262.2 | 12.1 | 21234 4 | US-09-810-671-3 Sequence 3, Appl |
| 18 | 262.2 | 12.1 | 21234 4 | US-10-109-854-3 Sequence 3, Appl |
| 19 | 260.4 | 12.0 | 63000 4 | US-09-780-172-18 Sequence 18, Appl |
| 20 | 259.2 | 11.9 | 3885 1 | US-08-688-145-1 Sequence 1, Appl |
| 21 | 258.6 | 11.9 | 43950 4 | US-09-735-934A-3 Sequence 3, Appl |
| 22 | 258.6 | 11.9 | 43950 4 | US-10-060-332-3 Sequence 3, Appl |
| 23 | 257.6 | 11.9 | 99500 4 | US-09-798-096-10 Sequence 10, Appl |
| 24 | 257.2 | 11.8 | 10684 3 | US-08-618-100B-3 Sequence 3, Appl |
| 25 | 256.8 | 11.8 | 8174 1 | US-07-914-281-5 Sequence 5, Appl |
| 26 | 256.8 | 11.8 | 8174 1 | US-08-393-246-5 Sequence 5, Appl |
| 27 | 256.8 | 11.8 | 8174 1 | US-08-525-058A-5 Sequence 5, Appl |

| | | | | |
|------|-------|------|----------|-------------------------------------|
| C 28 | 256.8 | 11.8 | 8174 2 | US-08-696-731-5 Sequence 5, Appl |
| C 29 | 256.8 | 11.8 | 8174 3 | US-09-042-531-5 Sequence 5, Appl |
| C 30 | 256.8 | 11.8 | 8174 5 | PCT-US91-00899-3 Sequence 3, Appl |
| C 31 | 256.8 | 11.8 | 29629 4 | US-09-729-995-3 Sequence 3, Appl |
| C 32 | 256.8 | 11.8 | 29629 4 | US-10-135-689-3 Sequence 3, Appl |
| C 33 | 254.8 | 11.7 | 118067 4 | US-09-497-855A-32 Sequence 32, Appl |
| C 34 | 254.4 | 11.7 | 98844 4 | US-09-791-211-10 Sequence 10, Appl |
| C 35 | 253.8 | 11.7 | 16063 4 | US-09-801-052-3 Sequence 3, Appl |
| C 36 | 253.8 | 11.7 | 16063 4 | US-10-020-121-3 Sequence 3, Appl |
| C 37 | 253.2 | 11.7 | 14581 4 | US-08-520-373D-4 Sequence 4, Appl |
| C 38 | 253.2 | 11.7 | 22481 4 | US-08-367-841A-43 Sequence 43, Appl |
| C 39 | 253.2 | 11.7 | 22481 5 | PCT-US95-07201-43 Sequence 43, Appl |
| C 40 | 253.2 | 11.7 | 22484 4 | US-09-875-223-2 Sequence 2, Appl |
| C 41 | 253.2 | 11.7 | 22484 4 | US-09-875-114-2 Sequence 2, Appl |
| C 42 | 252.2 | 11.6 | 7210 2 | US-08-257-963B-10 Sequence 10, Appl |
| C 43 | 252.2 | 11.6 | 7210 4 | US-08-367-841A-10 Sequence 10, Appl |
| C 44 | 252.2 | 11.6 | 7210 5 | PCT-US95-07201-10 Sequence 10, Appl |
| C 45 | 252.2 | 11.6 | 14581 4 | US-08-520-373D-4 Sequence 4, Appl |

ALIGNMENTS

RESULT 1
US-09-817-180-3
Sequence 3, Application US/09817180
Patent No. 6340584
GENERAL INFORMATION:
APPLICANT: GAN, Weiniu et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001183
CURRENT APPLICATION NUMBER: US/09/817,180
CURRENT FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 15297
TYPE: DNA
ORGANISM: Human
US-09-817-180-3

Query Match 12.9%; Score 280; DB 4; Length 15297;
Best Local Similarity 74.9%; Pred. No. 9.3e-55;
Matches 370; Conservative 0; Mismatches 110; Indels 14; Gaps 1;

| | | | |
|----|-------|--|-------|
| QY | 1019 | ATTACTATATTAAGAGAGTGTGACAAAACAGATGATTAAGCTGCAACCTGGCACAC | 1078 |
| DB | 11651 | ACTCCCTTAATCCCAACCTTCCACACAGCAGAAATAGATTAACCTGGCCAGTTCTCAC | 11710 |
| QY | 1079 | GGTCATAGTTCAGCTGCTGGAGGTGAGAGGAGAGAGTGAACACAGGTTTC | 1138 |
| DB | 11711 | GGCTGATCCAGCACTTGGAGGCTGAGTGGTGTATCATCTTGAGCCACGAGATTC | 11770 |
| QY | 1139 | AAGGCAGCTGGGCAACATTAACAAGATCTTCTCTCAAAAAAAAAAAAAAAAAAGA | 1198 |
| DB | 11771 | AAGATAGCTTGACACACATGAAATCCATCTGTCAAAAAAAAAATA----- | 11825 |
| QY | 1199 | AAGAGAGAGGCGCGGCTGGTGTCTCAACCTGTATCCAGCATTTGGAGGCCGAG | 1258 |
| DB | 11826 | -----GACTGGGACAGGCGGTCTCAACCTGTATCCAGCATTTGGAGGCCGAG | 11876 |
| QY | 1259 | CGGGGAGATACCTGTGTGTGAGGTTGAGACAGCTGGCCAACTGGAAACCC | 1318 |
| DB | 11877 | GAGGTGATACCTGTGTGTGAGGTTGAGACAGCAACATGTGAACCC | 11936 |
| QY | 1319 | GTCTGACTCAAAATGCAAAATTTAGCAGGCGGTGTAGCAGGCACTGTATCCAGCT | 1378 |
| DB | 11937 | ATCTCTACTAAATAATCAAAATTTAGCAGGATGTGTGACAGTCTGTATCCAGCT | 11996 |
| QY | 1379 | ACTTGGAGGCTGAGGACAGAGATCGCTTGAACCCAGAGGTGAGGTTCAGTAACT | 1438 |

Db 11997 ACTGGAGAGCTGAGGTGGGAGAAATTGTTGAACCCAGAGGCGAGGCTGCAGTGAGCC 12056
QY 1439 GAGATCGTGGCGTTGCACTCCAGCTGGGCGACAAGACAGACTGTCTCAGAAAAAA 1498
Db 12057 GAGATGTGGCACTGCACTCCAGCTGGGCGACAAGAGTGAATCTCATCTCAAAAAAA 12116
QY 1499 AAAAAAAAAAGAGA 1512
Db 12117 CCAAAAAACAAAA 12130

RESULT 2

US-10-003-295-3
; Sequence 3, Application US/10003295
; Patent No. 6686187
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001183DIV
; CURRENT APPLICATION NUMBER: US/10/003,295
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-003-295-3

Query Match 12.9%; Score 280; DB 4; Length 15297;
Best Local Similarity 74.9%; Pred. No. 9,3e-55;
Matches 370; Conservative 0; Mismatches 110; Indels 14; Gaps 1;

QY 1019 ATTACTATATAGAGAGTGTGACAAACAGATGATTAAGTGCAGACCGTGGCACAC 1078
Db 11651 ACTCCCTTAATGCCAAGCTTCCCAACAGGCAAGATTAAGATTAACCTGGCCAGTGTCTCAC 11710
QY 1079 GCTCATAGTTTCACTGCTGGTGGAGGTTGAGAGGAGGATGCTTGAACAAGAGTTC 1138
Db 11711 GCCTGATCCAGACACTTGGGAGGCTGAGCTGGGTGATCACTTGAGCCCAAGAGTTC 11770
QY 1139 AAGCCAGCTTGGGCAACATAAGATCTGTCTCTCAAAAAAAGAAAAAAGAAAAA 1198
Db 11771 AAGATCAGCTTGGACAAACAGTGAACCTCTGTCAAAAAAATACAAAAATA----- 11825
QY 1199 AAGAGAGAGGCGCGGCGTGTGCTCAAGCTGTATCCAGACCTTGGAGGCGAG 1258
Db 11826 -----GACTGGGCAAGGCTGCTCAAGCTGTATCCAGACCTTGGAGGCGAG 11876
QY 1259 CCGGCGGATCACTGTGTGTCAGAGTTGAGACCAAGCTGGCCAAATGGCAAAACCC 1318
Db 11877 GAGGTGATCACTGTGTGTCAGAGTTGAGACCAAGCAATGTGAACCC 11936
QY 1319 GTCTGACTCAAAATGCAAAATTAAGCCAGGCTGTGACAGGCACTGTATCCAGCT 1378
Db 11937 ATCTCTAATAAAATTAACAAATTAAGCCAGGATGTGGCACTGTATCCAGCT 11996
QY 1379 ACTTGGAGGCTGAGGCAAGAAATCCGTGAACCCAGAGGTGAGGTTGCAAGTAAGCT 1438
Db 11997 ACTTGGAGGCTGAGGTTGGAATTTGTAACCCAGAGGCTGAGAGGCC 12056
QY 1439 GAGATGTGGCTGCACTCCAGCTGGGCGACAAGACAGACTGTCTCAGAAAAAA 1498
Db 12057 GAGATGTGGCTGCACTCCAGCTGGGCGACAAGAGTGAATCTCTCAAAAAAA 12116
QY 1499 AAAAAAAAAAGAGA 1512
Db 12117 CCAAAAAACAAAA 12130

RESULT 3

US-09-078-294-7/c
; Sequence 7, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sert, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 11811
; TYPE: DNA
; ORGANISM: BAC-P2 contig 3
US-09-078-294-7

Query Match 12.6%; Score 274.2; DB 3; Length 11811;
Best Local Similarity 79.7%; Pred. No. 1.8e-53;
Matches 350; Conservative 0; Mismatches 83; Indels 6; Gaps 2;

QY 1070 GTGGCACACGCTCATAGTTCTAGCTGCTTGGAGGTTGAGAGGAGATGCTTGAACA 1129
Db 6663 GTGGCTCATGCTGTATATCCAGCACTTGGGAGGCGAGGTGGAGATGACTTGAGCT 6604
QY 1130 CAGGTGTTAAGCCAGCCTGGGCAACATTAAGATCTGTCTCAAAAAA 1189
Db 6603 CAGAGTTCAAGACCAAGCCTGGGCAACAGTGAAGCTCATCTTCAAAATACATTTA 6544
QY 1190 AAAAAAAGAAAGAGAGGCGCGGCTGTGTGCTCAAGCTGTATATCCAGCACTTTGG 1249
Db 6543 AAGGTAG-----CTGGGGGCAAGGTGTGTGCTGCTGATATCCAGCCTTTGG 6489
QY 1250 GAGGCGAGCGCGGCGGATCACTGTGTGACAGATTGAGACCAAGCTTGGCAACATGG 1309
Db 6488 GAGGTCAAGTGGGTGATCACTGAGGTCAAGATTGAGGCCAGCTGGCAACATGG 6429
QY 1310 CAAACCCGCTGTATCTAAATATG-AAAAATTAGCCAGGCTGTGACAGGCACTGT 1368
Db 6428 TGAATCCCATCTCTAATAAAATCAAAAAATTGACAGGCTGTGTGCGGCGCTGT 6369
QY 1369 AATCCAGCTACTTGGAGGCTGAGGCAAGATCGCTGAACCCAGAGGTTGAGGTT 1428
Db 6368 AATCCAGCTACTGAGAGGCTGAGGCAAGAAATCATTTGAACCCAGAGAGGTT 6309
QY 1429 GCAATAGCTAGATGTGCGCTTGGCACTCCAGCTGGGCGACAAGACAGACTGTGT 1488
Db 6308 GCAATAGCAAGATATGCAATTTGCACTCCAGCTGGCAAGAGCAAACTCTCA 6249
QY 1489 TCAGAAAAA 1507
Db 6248 AAAAAAAAAAAAAA 6230

RESULT 4

US-09-735-934A-3/c
; Sequence 3, Application US/09735934A
; Patent No. 6372468
; GENERAL INFORMATION:
; APPLICANT: Li, Jiajin et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000851
; CURRENT APPLICATION NUMBER: US/09/735,934A
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 43950
; TYPE: DNA

ORGANISM: Homo sapiens
US-09-735-934A-3

Query Match 12.4%; Score 269.8; DB 4; Length 43950;

Best Local Similarity 77.9%; Pred. No. 3.1e-52; Matches 339; Conservative 0; Mismatches 92; Indels 4; Gaps 1;

QY 1063 GCGAACCTGTGGACACGCTCATAGTCTTCTGCTGGAGGTTAGAGAGGAGATGGC 1122
DB 17761 GGGCACAAGTGGCACAACCTGTATCCGAACTTTGGAGGCTAGGTGGAGATGGC 17702
QY 1123 TTGAACACAGGTGTTCAAGGCCAGCCTGGCAACATTAACAGATCCTGTCTCAAAAA 1182
DB 17701 TTGAGCGCAGAAATTCAAGACCACTGGGCAACATAGACAGACCCCTATCTTAAAAA 17642
QY 1183 AAAAAAAAAAAAAA---AAGAGAGAGGCGCGGCGTGTGCTCAGCCTGTATCC 1238
DB 17641 AAAAAAAAAAAAAAAGAGAGAAATATAGTCAAGCGCGGTGCTCATGCTGTATCC 17582
QY 1239 CAGCACTTTGGAGGCGGCGGCGGAGTCACTGTGTCAAGATTTGAGACAGCT 1238
DB 17581 CAGCACTTTGGAGGCGGCGGCGGAGTCACTGTGTGAGGTTGAGACAGCCT 17522
QY 1299 GCGCAACATGGCAAAACCCCTCTGTACTCAAAATGCAAAATTAAGCAGCGTGTAGC 1358
DB 17521 GACCAACATGGCAAAACCCCTCTGTACTCAAAATGCAAAATTAAGTGTGGCCTGGGC 17462
QY 1359 AGGCACTGTATATCCAGTCTTGGAGGCTGAGGAGAGATCGCTTGAACCCAGGA 1418
DB 17461 GGGTGCCTGTATATCCAGTCTTGGAGGCTGAGGAGAGATCGCTTGAACCCAGGA 17402
QY 1419 GGTGAGGTTCAGTAACTGAGATCGTCCCTTGCATCTCCAGCTGGCGACAGAGCA 1478
DB 17401 GACAGAGGTTCAGTAACTGAGATCGTCCCTTGCATCTCCAGCTGGCGACAGAGCG 17342
QY 1479 AGACTGTCTCAGA 1493
DB 17341 AAACGCCGTCTAAAA 17327

RESULT 5
US-10-060-332-3/C

Sequence 3, Application US/10060332
Patent No. 6528294
GENERAL INFORMATION:
APPLICANT: Li, Jiajin et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
FILE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: CL000851D1V
CURRENT APPLICATION NUMBER: US/10/060,332
CURRENT FILING DATE: 2002-02-01
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 43950
TYPE: DNA
ORGANISM: Homo sapiens
US-10-060-332-3

Query Match 12.4%; Score 269.8; DB 4; Length 43950;
Best Local Similarity 77.9%; Pred. No. 3.1e-52;
Matches 339; Conservative 0; Mismatches 92; Indels 4; Gaps 1;

QY 1063 GCGAACCTGTGGACACGCTCATAGTCTTCTGCTGGAGGTTAGAGAGGAGATGGC 1122
DB 17761 GGGCACAAGTGGCACAACCTGTATCCGAACTTTGGAGGCTAGGTGGAGATGGC 17702
QY 1123 TTGAACACAGGTGTTCAAGGCCAGCCTGGCAACATTAACAGATCCTGTCTCAAAAA 1182
DB 17701 TTGAGCGCAGAAATTCAAGACCACTGGGCAACATAGACAGACCCCTATCTTAAAAA 17642
QY 1183 AAAAAAAAAAAAAA---AAGAGAGAGGCGCGGCGTGTGCTCAGCCTGTATCC 1238

DB 17641 AAAAAAAAAAAAAAAGAGAGAAATATAGTCAAGCGCGGTGCTCATGCTGTATCC 17582

QY 1239 CAGCACTTTGGAGGCGGCGGCGGAGTCACTGTGTCAAGATTTGAGACAGCCT 1298

DB 17581 CAGCACTTTGGAGGCGGCGGCGGAGTCACTGTGTGAGGTTGAGACAGCCT 17522

QY 1299 GCGCAACATGGCAAAACCCCTCTGTACTCAAAATGCAAAATTAAGCAGCGTGTAGC 1358

DB 17521 GACCAACATGGCAAAACCCCTCTGTACTCAAAATGCAAAATTAAGTGTGGCCTGGGC 17462

QY 1359 AGGCACTGTATATCCAGTCTTGGAGGCTGAGGAGAGATCGCTTGAACCCAGGA 1418

DB 17461 GGGTGCCTGTATATCCAGTCTTGGAGGCTGAGGAGAGATCGCTTGAACCCAGGA 17402

QY 1419 GGTGAGGTTCAGTAACTGAGATCGTCCCTTGCATCTCCAGCTGGCGACAGAGCA 1478

DB 17401 GACAGAGGTTCAGTAACTGAGATCGTCCCTTGCATCTCCAGCTGGCGACAGAGCG 17342

QY 1479 AGACTGTCTCAGA 1493

DB 17341 AAACGCCGTCTAAAA 17327

RESULT 6
US-09-798-096-10/C

Sequence 10, Application US/09798096
Patent No. 639378
GENERAL INFORMATION:
APPLICANT: Donna T. Ward
TITLE OF INVENTION: ANTISENSE MODULATION OF REG2 EXPRESSION
FILE REFERENCE: KTS-0207
CURRENT APPLICATION NUMBER: US/09/798,096
CURRENT FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 89
SEQ ID NO 10
LENGTH: 99500
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-09-798-096-10

Query Match 12.4%; Score 269.2; DB 4; Length 99500;
Best Local Similarity 73.7%; Pred. No. 5.9e-52;
Matches 386; Conservative 0; Mismatches 103; Indels 35; Gaps 2;

QY 1039 GTGACAAAACAGATATTAAGTGGGACACCGTGGACACGCTCATAGTCTT 1098

DB 43878 GGGAAAAAATAAATCTTGGTGGAGTGTGTCTATGACTGTATCCAGCCCTTT 43819

QY 1099 -GGAGGTTGAGGAGGAGATGCTTGAACACAGGTGTTCAAGGCCAGCTGGGACCA 1157

DB 43818 GGGAGGCCAAGGACAGAGATGCTTAACTCAAGAGTTCAAGACCAAGCTGGGACCA 43755

QY 1158 TAACAAGATCTGTCTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGAGAGA----- 1206

DB 43758 CAGTGAAGACCTGTCTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGTCTTGAATCTTTTCT 43699

QY 1207 -----GGGCGGGGCTGTGTGCTCAAGCTGCTATATCCAGCA 1243

DB 43698 TTTTCTTTTCTTTTAAAGTGGGCTGGGACCGTGTCTCAAGCTGTATCCAGCA 43639

QY 1244 CTTTGGAGGCGGAGCGGCGGAGTCACTGTGTCAAGAGTTTGAACAGCTGGGCA 1303

DB 43638 CTTTGGAGGCGGAGAGAGGTATGATCACTGAGTCAAGAGTTTGAACAGCTGGGCA 43579

QY 1304 ACATGGCAAAACCCGCTGTCTCAAAATGCAAAATTAAGCAGCGTGTAGCAGGCA 1363

DB 43578 ATGTGTGAAGAACCCGCTTTTCTAATAATAAATAATTAAGCGGAGATGTGTGCTG 43519

QY 1364 CTTGTATCCAGCTTCTTGGAGGCTGAGGAGAGATCGCTTGAACCCAGAGGTGG 1423

Db 43518 TGTGTAATCCAGACTCGGAGAGGCTGAGGACGAGAAATGCTTGAACCTGGAGAGGTGG 43459
QY 1424 AGGTTCGAGTAAGCTGAGATGTCGCGTTCGACTCCAGCTCGGCGACAGAGAACT 1483
Db 43458 AGGTTCGAGTAAGCTGAGATGTCGCGTTCGACTCCAGCTCGGCGACAGAAATGAGACT 43399
QY 1484 CTGTCTCAGAAAAAAGAGAGAGAGAGAGAGAGAGAG 1527
Db 43398 TTGTCTCAGAAAAAAGAGAGAGAGAGAGAGAGAGAG 43355

RESULT 7

US-09-813-817-3/c
; Sequence 3, Application US/09813817
; Patent No. 6340583
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001178
; CURRENT APPLICATION NUMBER: US/09/813,817
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-813-817-3

Query Match 12.3%; Score 267.4; DB 4; Length 59065;
Best Local Similarity 78.2%; Pred. No. 1.2e-51;
Matches 351; Conservative 0; Mismatches 86; Indels 12; Gaps 2;

QY 1058 AAGCTGCGAACCCTGCGACACCTCATAGTTTACTGCTGGAGGTTGAGAGAGAGG 1117
Db 23369 AGGCCAGCATAGTGGCTCATGCTATATCCAGC-ATTGAGAGGCCAGAGAGAGG 23311
QY 1118 ATGGCTTGAACACAGGTGTTCAAGGCGAGCTGGCAACATPACAAATCCTGCTCA 1177
Db 23310 ACTGCTTGAACACAGGTGTTGAGACACCTGGTGAACAAACAAACCTTATCTTAC 23251
QY 1178 AAAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1237
Db 23250 TAAAAAATCATGCAATCA-----GGCAGTATGTTGCTCAACGCTGTATTC 23202
QY 1238 CCAGCATTTGGAGGCGAGCCGCGGAGATCACTGTGTCAGAGATTGAGACCAAGC 1297
Db 23201 CCAGCATTTGGAGGCGTGAAGCGGCGAGATCACTGAGGTTGAGAGTTGAGACTAACC 23142
QY 1298 TGGCCAACTGGCAAAACCCGCTGTGTAATCAAAATGCAAAATTAAGCAGGCGGTAG 1357
Db 23141 TGACCAACATGAGAAACCTGTCTTACTATAAAATPACAAATTAAGCAGGCGGTAG 23082
QY 1358 CAGGCACTGTATCCAGCTACTTGGAGGCTGAGAGCGAGAGAAATCGTTGACCCAGG 1417
Db 23081 CAGTCTGTATTCAGCTACTCGAGAGGCTGAGAGCGAGAGAAATCGTTGACCCAGG 23022
QY 1418 AGGTGAGGTTCAGTAAGCTGAGATGTCGCGTTCGACTCCAGCTGGCGACAAAGAGC 1477
Db 23021 AGCCAGAGGTTCGAGTGAAGCTGAGTTCGCAATTCATTCAGCCTGGCGACAAAGAGT 22962
QY 1478 AAGACTCTGTCTCAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1506
Db 22961 GAAACTCTGTCTCAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 22933

RESULT 8
US-09-978-197-3/c
; Sequence 3, Application US/09978197
; Patent No. 6403353
; GENERAL INFORMATION:

; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001178D1V
; CURRENT APPLICATION NUMBER: US/09/978,197
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 09/813,817
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-978-197-3

Query Match 12.3%; Score 267.4; DB 4; Length 59065;
Best Local Similarity 78.2%; Pred. No. 1.2e-51;
Matches 351; Conservative 0; Mismatches 86; Indels 12; Gaps 2;

QY 1058 AAGCTGCGAACCCTGCGACACCTCATAGTTTACTGCTGGAGGTTGAGAGAGAGG 1117
Db 23369 AGGCCAGCATAGTGGCTCATGCTATATCCAGC-ATTGAGAGGCCAGAGAGAGG 23311
QY 1118 ATGGCTTGAACACAGGTGTTCAAGGCGAGCTGGCAACATPACAAATCCTGCTCA 1177
Db 23310 ACTGCTTGAACACAGGTGTTGAGACACCTGGTGAACAAACAAACCTTATCTTAC 23251
QY 1178 AAAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1237
Db 23250 TAAAAAATCATGCAATCA-----GGCAGTATGTTGCTCAACGCTGTATTC 23202
QY 1238 CCAGCATTTGGAGGCGAGCCGCGGAGATCACTGTGTCAGAGATTGAGACCAAGC 1297
Db 23201 CCAGCATTTGGAGGCGTGAAGCGGCGAGATCACTGAGGTTGAGAGTTGAGACTAACC 23142
QY 1298 TGGCCAACTGGCAAAACCCGCTGTGTAATCAAAATGCAAAATTAAGCAGGCGGTAG 1357
Db 23141 TGACCAACATGAGAAACCTGTCTTACTATAAAATPACAAATTAAGCAGGCGGTAG 23082
QY 1358 CAGGCACTGTATCCAGCTACTTGGAGGCTGAGAGCGAGAGAAATCGTTGACCCAGG 1417
Db 23081 CAGTCTGTATTCAGCTACTCGAGAGGCTGAGAGCGAGAGAAATCGTTGACCCAGG 23022
QY 1418 AGGTGAGGTTCAGTAAGCTGAGATGTCGCGTTCGACTCCAGCTGGCGACAAAGAGC 1477
Db 23021 AGCCAGAGGTTCGAGTGAAGCTGAGTTCGCAATTCATTCAGCCTGGCGACAAAGAGT 22962
QY 1478 AAGACTCTGTCTCAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1506
Db 22961 GAAACTCTGTCTCAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 22933

RESULT 9

US-09-608-285A-8/c
; Sequence 8, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:

; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11

```

; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3409)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc feature
; LOCATION: (9214)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
US-09-608-2854-8

Query Match      12.2% Score 265.6; DB 4; Length 9365;
Best Local Similarity 79.0%; Pred. No. 1.6e-51;
Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

QY 1143 CCAGCGTGGGCAACATTAACAGATCTGTCTCTCAAAAAAAAAAAAAAAAAAGAG- 1201
DB 7802 CCAGCATGGGTGACAGAGCAAGACCTGTCTCAAAAAAAAAAAAAAAAAAG 7743
QY 1202 -----AGAGAGGCGCGGCGTGTGGCTCAAGCCTGTATCCAGCACTTTGGAGG 1253
DB 7742 AAMGAMCAAAAAAGCGTGGGTGTGGCTCAATCCCTTATTCVAGCACTTTGGAGG 7683
QY 1254 CCAGAGCGCGGCGGATCAGCTGTGTGTCAGAGATTGAGACCAAGCTGGCCAAATGGCAA 1313
DB 7682 CYGAGGYGGGTGATCACTGAGRTCAAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 7623
QY 1314 ACCCGCTGTACTCAAAATGC--AAAAATTAGCAAGCGGTGTAGAGCAAGCACTGTAT 1371
DB 7622 ACCCGCTGTACTCAAAATGC--AAAAATTAGCAAGCGGTGTAGAGCAAGCACTGTAT 7563
QY 1372 CCCAGCTACTTGGAGGCTGAGGAGAGATCGCTTGAACCCAGAGAGTGTGAGGTTGCA 1431
DB 7562 CCCAGCTACTTGGAGGCTGAGGAGAGATCGCTTGAACCCAGAGAGTGTGAGGTTGCA 7503
QY 1432 GTAAGCTAGATCGTGGCGTTGCACTCAGCGCTGGCGCAAGAGCAAGACTGTGTCTCA 1491
DB 7502 GTGAGCYGAGTTCRYGCAATGCACTCAGCGCTGGGMAACAAGAGCAAGACTCGTCTCA 7443
QY 1492 GAAAAAAAAAAAAA 1507
DB 7442 AAAAAATATATATAA 7427
```

```

RESULT 10
US-09-350-836B-8/c
; Sequence 8, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
```

```

; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/09/350,836B
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(288)
; NAME/KEY: exon
; LOCATION: (1281)..(1580)
; NAME/KEY: exon
; LOCATION: (1820)..(1855)
; NAME/KEY: exon
; LOCATION: (2467)..(2555)
; NAME/KEY: exon
; LOCATION: (2863)..(2942)
; NAME/KEY: exon
; LOCATION: (3889)..(3950)
; NAME/KEY: exon
; LOCATION: (4894)..(4995)
; NAME/KEY: exon
; LOCATION: (5847)..(5987)
; NAME/KEY: exon
; LOCATION: (6966)..(7138)
; NAME/KEY: exon
; LOCATION: (8556)..(9365)
; NAME/KEY: misc feature
; LOCATION: (9214)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; NAME/KEY: misc feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; NAME/KEY: misc feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
US-09-350-836B-8

Query Match      12.2% Score 265.6; DB 4; Length 9365;
Best Local Similarity 79.0%; Pred. No. 1.6e-51;
Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

QY 1143 CCAGCGTGGGCAACATTAACAGATCTGTCTCTCAAAAAAAAAAAAAAAAAAGAG- 1201
DB 7802 CCAGCATGGGTGACAGAGCAAGACCTGTCTCAAAAAAAAAAAAAAAAAAG 7743
QY 1202 -----AGAGAGGCGCGGCGTGTGGCTCAAGCCTGTATCCAGCACTTTGGAGG 1253
DB 7742 AAMGAMCAAAAAAGCGTGGGTGTGGCTCAATCCCTTATTCVAGCACTTTGGAGG 7683
QY 1254 CCAGAGCGCGGCGGATCAGCTGTGTGTCAGAGATTGAGACCAAGCTGGCCAAATGGCAA 1313
DB 7682 CYGAGGYGGGTGATCACTGAGRTCAAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 7623
QY 1314 ACCCGCTGTACTCAAAATGC--AAAAATTAGCAAGCGGTGTAGAGCAAGCACTGTAT 1371
```

Db 7622 ACCCGCTCTACTATAAAATCAAAAATTAATGCGCGGCGTGGGCGCGCTGTAAT 7563
QY 1372 CCCAGCTACTTGGAGGCTGAGGAGGAGATCGCTTGAACCCAGAGGTGAGAGTTGCA 1431
Db 7562 CCCAGCTACTTGGAGGCTGAGGAGGAGATCGCTTGAACCCAGAGGTGAGAGTTGCA 7503
QY 1432 GTAGCTGAGATCGTGGCGCTTGGCACTCCAGCCTGGGCGACAGGCAAGACTGTCTCA 1491
Db 7502 GTGAGCTGAGATCGTGGCGCTTGGCACTCCAGCCTGGGCGACAGGCAAGACTGTCTCA 7443
QY 1492 GAAAAAATATATAATAA 1507
Db 7442 AAAAAATATATAATAA 7427

RESULT 11
US-09-370-265-8/c
; Sequence 8, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 2811/35908
; CURRENT APPLICATION NUMBER: US/09/370,265
; EARLIER FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: PCT/US99/16180
; EARLIER FILING DATE: 1999-07-16
; EARLIER APPLICATION NUMBER: 09/350,836
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: 09/273,447
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/244,444
; EARLIER FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 09/122,449
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: 09/118,205
; EARLIER FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (72)..(287)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1280)..(1579)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1819)..(1854)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2466)..(2555)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2863)..(2940)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3887)..(3952)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4896)..(4994)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5846)..(5986)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6965)..(7138)

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8556)..(8639)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3409)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9214)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
US-09-370-265-8

Query Match 12.2%; Score 265.6; DB 4; Length 9365;
Best local similarity 79.0%; Pred. No. 1.6e-51;
Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

QY 1143 CCAGCTGGGCAACTTACAGATCTGTCTCAAAAAAAAAAAAAAAAAAG- 1201
Db 7802 CCAGCTGGGTGACAGAGCAAGCCCTGTCTCAAAAAAAAAAAAAAAAAAG 7743
QY 1202 -----AGAGAGGCGCGCGGTGTGCTCAGGCTGTATCCAGCACTTTGGAGG 1253
Db 7742 AAMGAMCAAAAAGCTRGGAGTGTGTCTCAATCTTATCTTACACTTTGGAGG 7683
QY 1254 CCGAGCGGGCGGATCACTGTGTGTCAGAGTTTGAACAGCTGCGCAATGCGCAA 1313
Db 7682 CCGAGCGGGGTGATCACTGTGTGTCAGAGTTTGAAMTGTGCGCAATGCGTGA 7623
QY 1314 ACCCGCTGTACTCAAAATGCG-AAAAATTGCCAGCGGTGTGTCAGGCACTGTAT 1371
Db 7622 ACCCGCTCTACTATAAAATCAAAAATTAATGCGCGGCGTGTGCGCGCTGTAT 7563
QY 1372 CCCAGCTACTTGGAGGCTGAGGAGGAGATCGCTTGAACCCAGAGGTGAGAGTTGCA 1431
Db 7562 CCCAGCTACTTGGAGGCTGAGGAGGAGATCGCTTGAACCCAGAGGTGAGAGTTGCA 7503
QY 1432 GTAGCTGAGATCGTGGCGCTTGGCACTCCAGCCTGGGCGACAGGCAAGACTGTCTCA 1491
Db 7502 GTGAGCTGAGATCGTGGCGCTTGGCACTCCAGCCTGGGCGACAGGCAAGACTGTCTCA 7443
QY 1492 GAAAAAATATATAATAA 1507
Db 7442 AAAAAATATATAATAA 7427

RESULT 12
US-09-557-800C-8/c
; Sequence 8, Application US/09557800C
; Patent No. 6476211
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; FILE REFERENCE: 28110/36457
; CURRENT APPLICATION NUMBER: US/09/557,800C
; EARLIER FILING DATE: 2000-04-25
; EARLIER APPLICATION NUMBER: 09/481,238
; EARLIER FILING DATE: 2000-01-11
; EARLIER APPLICATION NUMBER: 09/370,265
; EARLIER FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: PCT/US99/16180
; EARLIER FILING DATE: 1999-07-16

PRIOR APPLICATION NUMBER: 09/350836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/122449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/244444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/118,205
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 56
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 8
LENGTH: 9365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (3409)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
FEATURE:
NAME/KEY: misc feature
LOCATION: (9214)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
FEATURE:
NAME/KEY: misc feature
LOCATION: (9303)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
FEATURE:
NAME/KEY: misc feature
LOCATION: (9311)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
US-09-557-800C-8

Query Match 12.2%; Score 265.6; DB 4; Length 9365;
Best Local Similarity 79.0%; Pred. No. 1.6e-51;
Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

QY 1143 CCAGCTGGGCAATACAGATCCTCTCTCAAAAAAAAAAAAAAAAAAGAG- 1201
DB 7802 CCAGCATGGGTGACAGAGACCTGTCTCAAAAAAAAAAAAAAAAAAG 7743
QY 1202 -----AGAGGGGCGGCGGTGTGCTCAAGCTGTATCCAGCACTTTGGAGG 1253
DB 7742 AAWCGMCAAAAAAAGGCGGAGTGTGGCTCAATCTTAATTCAGCACTTTGGAGG 7683
QY 1254 CCAGCGCGGCGGATCACTGTGTGAGAGTTGAGACAGCTTGCCAACTGGCAA 1313
DB 7682 CYAGAGYGGGTGATCACTGAGRTCAGAGTTCAGAGTTCAGCACTTTGGAGG 7623
QY 1314 ACCCGCTGTACTCAAAATGC--AAAAATTAGCCAGGCGGTGTAGCAGCACTGTAT 1371
DB 7622 ACCCGCTGTACTTAATAATCAAAAAATTAGCCAGGCGGTGTAGCAGCACTGTAT 7563
QY 1372 CCCAGCTACTTGGAGGCTGAGCAGAGATGCTTGAACCCAGAGGTGAGGTGCA 1431
DB 7562 CCAGCTACTTGGAGGCTGAGCAGAGATGCTTGAACCCAGAGGTGAGGTGCA 7503
QY 1432 GTAAGCTGAGATCGTGGCGTTGCACTCCAGCTGGCGGCAAGAGCAATCTGTCTCA 1491
DB 7502 GTAGGCGWGTATRYGCAATGCACTCCAGCTGGGAAACAAGAGCAGACCTCGTCTCA 7443
QY 1492 GAAAAAAAAAAAAA 1507
DB 7442 AAAAAAAAAATTAATTA 7427

RESULT 13
US-09-370-625A-8/c
Sequence 8, Application US/09370625A
Patent No. 6600032
GENERAL INFORMATION:
APPLICANT: Ford, John

APPLICANT: Mulero, Julio
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
FILE REFERENCE: 28110/35808
CURRENT APPLICATION NUMBER: US/09/370,625A
CURRENT FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350,836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
PRIOR FILING DATE: 1999-03-19
NUMBER OF SEQ ID NOS: 39
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 8
LENGTH: 9365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: exon
LOCATION: (1)..(288)
FEATURE:
NAME/KEY: exon
LOCATION: (1281)..(1580)
FEATURE:
NAME/KEY: exon
LOCATION: (1820)..(1855)
FEATURE:
NAME/KEY: exon
LOCATION: (2863)..(2942)
FEATURE:
NAME/KEY: exon
LOCATION: (3889)..(3950)
FEATURE:
NAME/KEY: exon
LOCATION: (4894)..(4995)
FEATURE:
NAME/KEY: exon
LOCATION: (5847)..(5987)
FEATURE:
NAME/KEY: exon
LOCATION: (6966)..(7138)
FEATURE:
NAME/KEY: exon
LOCATION: (8556)..(9365)
FEATURE:
NAME/KEY: misc feature
LOCATION: (3409)
OTHER INFORMATION: n = a or g or t or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (9214)
OTHER INFORMATION: n = a or g or t or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (9303)
OTHER INFORMATION: n = a or g or t or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (9311)
OTHER INFORMATION: n = a or g or t or c
US-09-370-625A-8

Query Match 12.2%; Score 265.6; DB 4; Length 9365;
Best Local Similarity 79.0%; Pred. No. 1.6e-51;
Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

QY 1143 CCAGCTGGGCAATACAGATCCTCTCTCAAAAAAAAAAAAAAAAAAGAG- 1201
DB 7802 CCAGCATGGGTGACAGAGACCTGTCTCAAAAAAAAAAAAAAAAAAG 7743

QY 1202 -----AGAGAGGCGCGGCTGTGATCAAGCTGTATCCAGCACTTTGGAGG 1253
DB 7742 AAMCAGCAAAAAAGGCTGAGGTGTGCTCAATCTTATTCATCTTTGGAGG 7653
QY 1254 CCGAGCGGCGCGGATCACTGTGTGAGAGTTTGAACAAGCTGGCAATGGCAAA 1313
DB 7662 CCGAGGCGGCTGTGATCACTGTGTGAGAGTTTGAAGAGCTGGCAATGGTGA 7623
QY 1314 ACCCGCTGTCTCAAAAAATGC-AAAAATTGCGAGCGGTGTGAGCAAGCACTGTAT 1371
DB 7622 ACCCGCTGTCTCAAAAAATGC-AAAAATTGCGAGCGGTGTGAGCAAGCACTGTAT 7563
QY 1372 CCGAGCTGTGTGAGCGGTGTGAGCAAGCACTGTGAAACCGAGAGGTGAGGTGA 1431
DB 7562 CCGAGCTGTGTGAGCGGTGTGAGCAAGCACTGTGAAACCGAGAGGTGAGGTGA 7503
QY 1432 GTAAGCTGAGATCGTCCGTTGCACTCCAGCTGGCGCAAGAGCAAGCACTGTCTCA 1491
DB 7502 GTGAGCGGATCGTCCGTTGCACTCCAGCTGGCGCAAGAGCAAGCACTGTCTCA 7443
QY 1492 GAAAAAATATATATAA 1507
DB 7442 AAAAAAATATATATAA 7427

RESULT 14

US-09-608-285A-42/c
Sequence 42, Application US/09608285A
Patent No. 6335013
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
APPLICANT: Yeung, George
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
FILE REFERENCE: 28110/36570
CURRENT APPLICATION NUMBER: US/09/608,285A
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/583,231
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 09/557,800
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481,238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370,265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350,836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/244,444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/122,449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/118,205
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 42
LENGTH: 14747
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (13641)
OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine
US-09-608-285A-42

Query Match 12.2%; Score 265.6; DB 4; Length 14747;
Best Local Similarity 79.0%; Pred. No. 1.9e-51;
Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

QY 1143 CCGAGCTGGCAACATATACAGATCTCTCAAAAAAAAAAAAAAAAAAGAAAG- 1201
DB 1162 CCGAGCTGGGTGACAGAGCAACCTCTCTCAAAAAAAAAAAAAAAAAAG 1103
QY 1202 -----AGAGAGGCGCGGCTGTGATCAAGCTGTATCCAGCACTTTGGAGG 1253
DB 1102 AAMCAGCAAAAAAGGCTGAGGTGTGCTCAATCTTATTCATCTTTGGAGG 11043
QY 1254 CCGAGCGGCGCGGATCACTGTGTGAGAGTTTGAACAAGCTGGCAATGGCAAA 1313
DB 11042 CCGAGGCGGCTGTGATCACTGTGTGAGAGTTTGAAGAGCTGGCAATGGTGA 10983
QY 1314 ACCCGCTGTCTCAAAAAATGC-AAAAATTGCGAGCGGTGTGAGCAAGCACTGTAT 1371
DB 10982 ACCCGCTGTCTCAAAAAATGC-AAAAATTGCGAGCGGTGTGAGCAAGCACTGTAT 10923
QY 1372 CCGAGCTGTGTGAGCGGTGTGAGCAAGCACTGTGAAACCGAGAGGTGAGGTGA 1431
DB 10922 CCGAGCTGTGTGAGCGGTGTGAGCAAGCACTGTGAAACCGAGAGGTGAGGTGA 10863
QY 1432 GTAAGCTGAGATCGTCCGTTGCACTCCAGCTGGCGCAAGAGCAAGCACTGTCTCA 1491
DB 10862 GTGAGCGGATCGTCCGTTGCACTCCAGCTGGCGCAAGAGCAAGCACTGTCTCA 10803
QY 1492 GAAAAAATATATATAA 1507
DB 10802 AAAAAAATATATATAA 10787

RESULT 15

US-09-557-800C-42/c
Sequence 42, Application US/09557800C
Patent No. 6476211
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
APPLICANT: Yeung, George
TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
FILE REFERENCE: 28110/36457
CURRENT APPLICATION NUMBER: US/09/557,800C
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481,238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370,265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350,836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/244,444
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/122,449
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 09/118,205
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 56
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 42
LENGTH: 14747
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (13641)
OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine
US-09-557-800C-42

Query Match 12.2%; Score 265.6; DB 4; Length 14747;
Best Local Similarity 79.0%; Pred. No. 1.9e-51;
Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

```
QY 1143 CCAGCTGGGCAACATACAGATCTGTCTCTCAAAAAAAAAAAAAAAAAAGAAAG- 1201
    |||||
Db 11162 CCAGCATGGGTACAGACCAAGACCCTGTCTCAAAAAAAAAAAAAAAAAAG 11103
    |||||
QY 1202 -----AGAGAGGGCCGGGGGTGGTGGCTCAGCGCTGTATCCAGCACTTTGGGAGG 1253
    |||||
Db 11102 AAMCGAMCAAAAAAAAAAGGCTRGKVGTTGGCTCAIRCTRTATYCTAGCACTTTGGGAGG 11043
    |||||
QY 1254 CCGAGCCGGGGCGGATCACTGTGTGTCAGAGTTTGAACAGCCTGGCCACATGGCAAA 1313
    |||||
Db 11042 CYGAGGYGGGTGGATCACCTGAGRTCAGAGTTTGAAGMYAGCCTGGCCACATGGTGA 10983
    |||||
QY 1314 ACCCGTCTGTACTCAAAATGC--AAAAATTAGCCAGGCGTGTAGCAGGCACCTGTAT 1371
    |||||
Db 10982 ACCCGTCTGTACTCAAAATGC--AAAAATTAGCCAGGCGTGTAGCAGGCACCTGTAT 10923
    |||||
QY 1372 CCCAGCTACTTGGGAGGCTGAGGCGAGAAATCGCTTGAACCCAGGAGGTGAGGTTGA 1431
    |||||
Db 10922 CCCAGCTACTTGGGAGGCTGAGGCGAGAAATCGCTTGAACCCAGGAGGTGAGGTTGA 10863
    |||||
QY 1432 GTAAGCTGAGATCGTCCGTTGCACTCCAGCCTGGGCGACAAGACAAGACTGTCTCA 1491
    |||||
Db 10862 GTGAGCTGAGATCGTCCGTTGCACTCCAGCCTGGGCGACAAGACAAGACTGTCTCA 10803
    |||||
QY 1492 GAAAAAAAAAAAAAAAAA 1507
    |||||
Db 10802 AAAAAAAAAATAATAA 10787
    |||||
```

Search completed: March 13, 2004, 00:51:28
Job time : 166.607 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 21:22:51 ; Search time 762.198 Seconds

(without alignments)
10491.276 Million cell updates/sec

Title: US-09-966-880A-15

Perfect score: 2172
Sequence: 1 cccctgatacaggttgatga.....agtcgtatgataacttg 2172

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2432557 seqs, 184079884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEM_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 2172 | 100.0 | 2172 | 9 | US-09-966-880A-15 |
| 2 | 2172 | 100.0 | 2818 | 9 | US-09-966-880A-7 |
| 3 | 2172 | 100.0 | 6564 | 9 | US-09-966-880A-10 |
| 4 | 2172 | 100.0 | 11204 | 9 | US-09-966-880A-35 |
| 5 | 295.2 | 13.6 | 13862 | 10 | US-09-764-891-5477 |
| 6 | 295.2 | 13.6 | 13862 | 10 | US-09-764-891-10204 |
| 7 | 295.2 | 13.6 | 13862 | 14 | US-10-205-428-1003 |
| 8 | 292.6 | 13.5 | 7809 | 10 | US-09-764-891-6094 |
| 9 | 284.4 | 13.1 | 32249 | 10 | US-09-764-891-7364 |
| 10 | 281.4 | 13.0 | 65608 | 9 | US-09-962-436-292 |
| 11 | 281.4 | 13.0 | 65608 | 9 | US-09-962-832-119 |
| 12 | 281.4 | 13.0 | 65608 | 9 | US-09-954-531-180 |
| 13 | 280.2 | 12.9 | 11821 | 9 | US-09-764-877-2857 |
| 14 | 280.2 | 12.9 | 11821 | 15 | US-10-242-515-2857 |
| 15 | 280 | 12.9 | 15297 | 13 | US-10-003-295-3 |

| | | | | | | |
|----|-------|------|--------|----|----------------------|--------------------|
| 16 | 278.2 | 12.8 | 32204 | 10 | US-09-764-891-8537 | Sequence 8537, App |
| 17 | 277.4 | 12.8 | 30000 | 14 | US-10-262-552-33 | Sequence 33, Appl |
| 18 | 276.6 | 12.7 | 37940 | 15 | US-10-348-072A-12 | Sequence 12, Appl |
| 19 | 276.6 | 12.7 | 38000 | 15 | US-10-175-627-11 | Sequence 11, Appl |
| 20 | 276.4 | 12.7 | 176080 | 12 | US-10-235-192A-43 | Sequence 43, Appl |
| 21 | 275.6 | 12.7 | 160361 | 12 | US-10-235-192A-35 | Sequence 35, Appl |
| 22 | 275.4 | 12.7 | 3966 | 9 | US-09-764-855-227 | Sequence 227, App |
| 23 | 275.4 | 12.7 | 3966 | 14 | US-10-072-349-227 | Sequence 227, App |
| 24 | 275 | 12.7 | 6482 | 14 | US-10-180-410-6 | Sequence 6, Appl |
| 25 | 275 | 12.7 | 20050 | 14 | US-10-180-410-5 | Sequence 5, Appl |
| 26 | 274.8 | 12.7 | 15362 | 9 | US-09-764-877-2856 | Sequence 2856, Ap |
| 27 | 274.8 | 12.7 | 15362 | 15 | US-10-242-515-2856 | Sequence 2856, Ap |
| 28 | 274.2 | 12.6 | 11811 | 10 | US-09-128-552-7 | Sequence 7, Appl |
| 29 | 274.2 | 12.6 | 176001 | 16 | US-10-210-556-27 | Sequence 27, Appl |
| 30 | 274.2 | 12.6 | 186739 | 16 | US-10-210-556-19 | Sequence 43, Appl |
| 31 | 274 | 12.6 | 4963 | 9 | US-09-764-877-2903 | Sequence 2903, Ap |
| 32 | 274 | 12.6 | 4963 | 15 | US-10-242-515-2903 | Sequence 2903, Ap |
| 33 | 273.6 | 12.6 | 174424 | 9 | US-09-967-768A-314 | Sequence 314, App |
| 34 | 273.6 | 12.6 | 174424 | 10 | US-09-960-706-969 | Sequence 969, App |
| 35 | 273 | 12.6 | 2401 | 15 | US-10-027-632-102635 | Sequence 102635, |
| 36 | 273 | 12.6 | 26657 | 9 | US-09-810-673A-3 | Sequence 3, Appl |
| 37 | 273 | 12.6 | 26657 | 14 | US-10-395-242-3 | Sequence 3, Appl |
| 38 | 272 | 12.5 | 17491 | 14 | US-10-017-161-1995 | Sequence 1995, Ap |
| 39 | 270.8 | 12.5 | 670 | 15 | US-10-027-632-201375 | Sequence 201375, |
| 40 | 270.8 | 12.5 | 1061 | 14 | US-10-277-161-66 | Sequence 66, Appl |
| 41 | 270.8 | 12.5 | 107880 | 10 | US-09-792-616-1 | Sequence 1, Appl |
| 42 | 270.2 | 12.4 | 17286 | 9 | US-09-764-877-3234 | Sequence 3234, Ap |
| 43 | 270.2 | 12.4 | 17286 | 15 | US-10-242-515-3234 | Sequence 3234, Ap |
| 44 | 269.8 | 12.4 | 43950 | 13 | US-10-060-332-3 | Sequence 3, Appl |
| 45 | 269.8 | 12.4 | 43950 | 14 | US-10-339-657-3 | Sequence 3, Appl |

ALIGNMENTS

RESULT 1
US-09-966-880A-15
Sequence 15, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Yasuku
APPLICANT: Muramatsu, Masamichi
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966,880A
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 2172
TYPE: DNA
ORGANISM: Homo sapiens
US-09-966-880A-15
Query Match 100.0%; Score 2172; DB 9; Length 2172;
Best local similarity 100.0%; Pred. No. 0;
Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CCCCTGATAGGTTGATGACTTACGAGAGCATTTGCGATCTTGGACCTTATGCGCA 60
DB 1 CCCCTGATAGGTTGATGACTTACGAGAGCATTTGCGATCTTGGACCTTATGCGCA 60
OY 61 CTTCCAGAGATGTCACACAGATGATATCTCTGCTGAGAGACAGTGAATTAACAGT 120
DB 61 CTTCCAGAGATGTCACACAGATGATATCTCTGCTGAGAGACAGTGAATTAACAGT 120

QY 121 CCTCAAGTCTTCTGTGTTTATTTCTTCAACTCTCACTTTCTAGAGTTTACAGAAAA 180
DB 121 CCTCAAGTCTTCTGTGTTTATTTCTTCAACTCTCACTTTCTAGAGTTTACAGAAAA 180
QY 181 AATATTATATAGACTCTTTAAAAAGATCTATGCTTGAAGAAATAGAGAAAGAAACAGCT 240
DB 181 AATATTATATAGACTCTTTAAAAAGATCTATGCTTGAAGAAATAGAGAAAGAAACAGCT 240
QY 241 CTGGCCAGGAGCGTGTCTCAATTTGGTGCAGTTTGAATGCAACTTGTCCCTCACTGGA 300
DB 241 CTGGCCAGGAGCGTGTCTCAATTTGGTGCAGTTTGAATGCAACTTGTCCCTCACTGGA 300
QY 301 ATTAACAGAACTGACGAGACCTGGAGACATCTTAAAGTGTCAAGCTTTTCTATGACTTTTA 360
DB 301 ATTAACAGAACTGACGAGACCTGGAGACATCTTAAAGTGTCAAGCTTTTCTATGACTTTTA 360
QY 361 GGTAGAGATGAGAGAGAGAGTATCTTAAAGATGATGAGAGATCAAAATGTTTTTA 420
DB 361 GGTAGAGATGAGAGAGAGAGTATCTTAAAGATGATGAGAGATCAAAATGTTTTTA 420
QY 421 TATCAACATCTTATATTTATTTGATTCATTTGATTTGATTTGATTTGATTTGATTTT 480
DB 421 TATCAACATCTTATATTTATTTGATTTGATTTGATTTGATTTGATTTGATTTT 480
QY 481 TCTATTCCTTTCCCTTGAAGTTTACTTTCAAGTAAACAAACTCTTCATCAGCCATGA 540
DB 481 TCTATTCCTTTCCCTTGAAGTTTACTTTCAAGTAAACAAACTCTTCATCAGCCATGA 540
QY 541 TCTATGAGACCTCTTATGAGAGTATCTGGGTGATTTGAGACCCCAACATCTCTCCAA 600
DB 541 TCTATGAGACCTCTTATGAGAGTATCTGGGTGATTTGAGACCCCAACATCTCTCCAA 600
QY 601 GATTAATATCCAAATCATGCGCTGATGTTTAAATCAGAGAGAGATGTTTATGTTG 660
DB 601 GATTAATATCCAAATCATGCGCTGATGTTTAAATCAGAGAGAGATGTTTATGTTG 660
QY 661 TACAAAAGAGATTTGATGAGTGGAGTGAAGTATGAGATGAGATGATGATGATGATGATG 720
DB 661 TACAAAAGAGATTTGATGAGTGGAGTGAAGTATGAGATGAGATGATGATGATGATG 720
QY 721 GCTACTTATATAAGAGATCTTAAATGAGAGAGAGATGATGATGATGATGATGATG 780
DB 721 GCTACTTATATAAGAGATCTTAAATGAGAGAGAGATGATGATGATGATGATGATG 780
QY 781 TGGGTGATGCTGAAGTGAAGTATCTTGAAGAGCAACTCTTTAAGAGATGCTCT 840
DB 781 TGGGTGATGCTGAAGTGAAGTATCTTGAAGAGCAACTCTTTAAGAGATGCTCT 840
QY 841 AATTGAGAAACCCCAAACTTCAATATCATATTTAGCAAAACATTTGAGAGAGATG 900
DB 841 AATTGAGAAACCCCAAACTTCAATATCATATTTAGCAAAACATTTGAGAGAGATG 900
QY 901 CTGGAATGTTGGGAGAGAGAAATCTATTGGCTCTGAGGCTCTCTCACTCAGAAATG 960
DB 901 CTGGAATGTTGGGAGAGAGAAATCTATTGGCTCTGAGGCTCTCTCACTCAGAAATG 960
QY 961 CCAATCAGGTCAAGTTTGTCTATTTTGTATGTTGTGTGTGTGTGTGTGTGTGTGTGT 1020
DB 961 CCAATCAGGTCAAGTTTGTCTATTTTGTATGTTGTGTGTGTGTGTGTGTGTGTGTGT 1020
QY 1021 TAACTATATAGAGATGATGAGAGAGATGATGAGAGATGATGAGAGATGATGAGAGAT 1080
DB 1021 TAACTATATAGAGATGATGAGAGAGATGATGAGAGATGATGAGAGATGATGAGAGAT 1080
QY 1081 TCAATAGTTAGCTGCTTGGAGGTTGAGAGAGAGATGATGAGAGATGATGAGAGATG 1140
DB 1081 TCAATAGTTAGCTGCTTGGAGGTTGAGAGAGAGATGATGAGAGATGATGAGAGATG 1140
QY 1141 GGCAGAGCTGGGCAACATTAAGATCTCTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1200
DB 1141 GGCAGAGCTGGGCAACATTAAGATCTCTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1200

QY 1201 GAGAGAGGCGGGGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 GAGAGAGGCGGGGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 GGGCGGATCACTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1261 GGGCGGATCACTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 CTGTACTGAATGAGAAATTAAGCAGGCGGTGATGATGATGATGATGATGATGATGATG 1380
DB 1321 CTGTACTGAATGAGAAATTAAGCAGGCGGTGATGATGATGATGATGATGATGATGATG 1380
QY 1381 TTGGAGGCTGAGAGAGAGATCGCTTGAACCCAGAGAGTGTGAGAGTGTGAGAGTGTG 1440
DB 1381 TTGGAGGCTGAGAGAGAGATCGCTTGAACCCAGAGAGTGTGAGAGTGTGAGAGTGTG 1440
QY 1441 GATCGTCCGTTGACTCCAGCTGGGCGACAGAGAGTGTGATGATGATGATGATGATGAT 1500
DB 1441 GATCGTCCGTTGACTCCAGCTGGGCGACAGAGAGTGTGATGATGATGATGATGATGAT 1500
QY 1501 AAAAAAAG 1560
DB 1501 AAAAAAAG 1560
QY 1561 TGCAGAGAAATGCTTATATCAACAAATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB 1561 TGCAGAGAAATGCTTATATCAACAAATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1621 TCTCTTTGCTGTCTATTTGCTTCACTGCTTGTGATGATGATGATGATGATGATGATG 1680
DB 1621 TCTCTTTGCTGTCTATTTGCTTCACTGCTTGTGATGATGATGATGATGATGATGATG 1680
QY 1681 ATAAACATATCCCTGTGCGGTATTTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
DB 1681 ATAAACATATCCCTGTGCGGTATTTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
QY 1741 ACAGAGAACTGAAAG 1800
DB 1741 ACAGAGAACTGAAAG 1800
QY 1801 AGAGTTAAAAATGTTACTTCAATGATGATGATGATGATGATGATGATGATGATGATG 1860
DB 1801 AGAGTTAAAAATGTTACTTCAATGATGATGATGATGATGATGATGATGATGATGATG 1860
QY 1861 AATTTTATTAACATGATTTCTTTCTGATATATGATGATGATGATGATGATGATGATG 1920
DB 1861 AATTTTATTAACATGATTTCTTTCTGATATATGATGATGATGATGATGATGATGATG 1920
QY 1921 AATTATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
DB 1921 AATTATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
QY 1981 GGTGCTAGAGAGAGATTTCTGATTTTATGATGATGATGATGATGATGATGATGATGAT 2040
DB 1981 GGTGCTAGAGAGAGATTTCTGATTTTATGATGATGATGATGATGATGATGATGATGAT 2040
QY 2041 TGGCTCACTTCAATCAGTTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
DB 2041 TGGCTCACTTCAATCAGTTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
QY 2101 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2160
DB 2101 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2160
QY 2161 GGAATTAACCTG 2172
DB 2161 GGAATTAACCTG 2172

GENERAL INFORMATION:

APPLICANT: Honjo, Tasuku
APPLICANT: Muramatsu, Masamichi
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966,880A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 2818
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (80)...(673)
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)...(79)
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (677)...(2818)
US-09-966-880A-7

Query Match 100.0%; Score 2172; DB 9; Length 2818;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCTGTATGAGTGTGATGACTTACGAGACGATTCCTTCTTGGACCTTTGATAGCA 60
DB 623 CCCCTGTATGAGTGTGATGACTTACGAGACGATTCCTTCTTGGACCTTTGATAGCA 682
QY 61 CTTCCAGGATGTCAACACGATGAAATATCTTGTCTGAAGACAGTGTATAAAAACGT 120
DB 683 CTTCCAGGATGTCAACACGATGAAATATCTTGTCTGAAGACAGTGTATAAAAACGT 742
QY 121 CCTTAAGTCTCTGTTTTTATCTCAACTCTCACTTCTTGAAGTTACAGAAAA 180
DB 743 CCTTAAGTCTCTGTTTTTATCTCAACTCTCACTTCTTGAAGTTACAGAAAA 802
QY 181 AATATTTATACGACTCTTTAAAAAGATCTATGCTTGAATAAGAGAGAACACAGGT 240
DB 803 AATATTTATACGACTCTTTAAAAAGATCTATGCTTGAATAAGAGAGAACACAGGT 862
QY 241 CTGGCCAGGAGCGTCTGCAATGTGTGCAATTTGAATGCAACATTTGCCCTACTGGA 300
DB 863 CTGGCCAGGAGCGTCTGCAATGTGTGCAATTTGAATGCAACATTTGCCCTACTGGA 922
QY 301 ATPAACGAATGCAAGGACCTGGAGACATCTAAAGTGTCAAGTTTTCTATGACTTTTA 360
DB 923 ATPAACGAATGCAAGGACCTGGAGACATCTAAAGTGTCAAGTTTTCTATGACTTTTA 982
QY 361 GGTAGGATGAGAGAGAGATGATCTTAAAAAGATGGTGAAGATCAATGTTTTTA 420
DB 983 GGTAGGATGAGAGAGAGATGATCTTAAAAAGATGGTGAAGATCAATGTTTTTA 1042
QY 421 TATCAACATCTTTATTTATTTGATTCATTTGATTAACAGTGTGTATGATAGATTT 480
DB 1043 TATCAACATCTTTATTTATTTGATTCATTTGATTAACAGTGTGTATGATAGATTT 1102
QY 481 TGTATTTCTTTTCTTGAAGTTTACTTTAAGTAAACAAATCTTCCATCAGGCGCAGA 540
DB 1103 TGTATTTCTTTTCTTGAAGTTTACTTTAAGTAAACAAATCTTCCATCAGGCGCAGA 1162
QY 541 TCTATAGACCTCTTAATGAGATCTGTGGTGTGATTTGACCCCAACCATCTCTCCAA 600

DB 1163 TCTATAGACCTCTTAATGAGATCTGTGGTGTGATTTGACCCCAACCATCTCTCCAA 1222
QY 601 GCATTATATCCATATATGCGTGTATGTTTTTAATCAGAGAAAGATTTTTATGTTTG 660
DB 1223 GCATTATATCCATATATGCGTGTATGTTTTTAATCAGAGAAAGATTTTTATGTTTG 1282
QY 661 TCAAAAAGAAATGTTATGGGTGGGAGTGAAGTATAGACATGATGATGATGATGAT 720
DB 1283 TCAAAAAGAAATGTTATGGGTGGGAGTGAAGTATAGACATGATGATGATGATGAT 1342
QY 721 GCTACTTTATTAAGATCTTAAAAATGGGAGAGAGCTGTGAACAGACACCTTAATA 780
DB 1343 GCTACTTTATTAAGATCTTAAAAATGGGAGAGAGCTGTGAACAGACACCTTAATA 1402
QY 781 TGGGTGATGTCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 840
DB 1403 TGGGTGATGTCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1462
QY 841 AATTTGAAACACCCCAACCTTCAATATCAATTAATGCAACAAATGGAAGAGTTG 900
DB 1463 AATTTGAAACACCCCAACCTTCAATATCAATTAATGCAACAAATGGAAGAGTTG 1522
QY 901 CTTGAATGTGGGAGAGAGAAATCTATTGCTCTGCTGCTCTCTCAATCTCAGAAATG 960
DB 1523 CTTGAATGTGGGAGAGAGAAATCTATTGCTCTGCTGCTCTCTCAATCTCAGAAATG 1582
QY 961 CCATCAGGTCAAGGTTTCTCAATTTTGTATGTGTGTATGATCTTCCCAAGTATAT 1020
DB 1583 CCATCAGGTCAAGGTTTCTCAATTTTGTATGTGTGTATGATCTTCCCAAGTATAT 1642
QY 1021 TAACTATATAGAGAGTGTGACAAAACGAATGATTAAGTGTGCAACCTGTGCAACGC 1080
DB 1643 TAACTATATAGAGAGTGTGACAAAACGAATGATTAAGTGTGCAACCTGTGCAACGC 1702
QY 1081 TCAATTTCTAGTGTGTTGGAGTGTGAAGAGAGAGATGCTTGAACACAGTGTGCA 1140
DB 1703 TCAATTTCTAGTGTGTTGGAGTGTGAAGAGAGAGATGCTTGAACACAGTGTGCA 1762
QY 1141 GGCCAGCTGGGCAATATCAAGATCTGTCTCTCAAAAAAAGAAAAAAGAAA 1200
DB 1763 GGCCAGCTGGGCAATATCAAGATCTGTCTCTCAAAAAAAGAAAAAAGAAA 1822
QY 1201 GAGAGAGGCGCGGCGTGTGTGCTCAGCCTGTAAATCCAGACATTTGGAGGCGGAGCC 1260
DB 1823 GAGAGAGGCGCGGCGTGTGTGCTCAGCCTGTAAATCCAGACATTTGGAGGCGGAGCC 1882
QY 1261 GGGCGGATCACTGTGTGTGAGAGATTGAGACCAAGCTGTGCAACATGGCAAAACCCCGT 1320
DB 1883 GGGCGGATCACTGTGTGTGAGAGATTGAGACCAAGCTGTGCAACATGGCAAAACCCCGT 1942
QY 1321 CTGTACTCAAAAAGCAAAAATTAAGCAGGCGGTGTAGAGGCACTGTAAATCCAGCTAC 1380
DB 1943 CTGTACTCAAAAAGCAAAAATTAAGCAGGCGGTGTAGAGGCACTGTAAATCCAGCTAC 2002
QY 1381 TTGGAGGCTGAGGAGAGAAATCGCTTGAACCCAGAGTGTGAGGTTGCAATGACTGA 1440
DB 2003 TTGGAGGCTGAGGAGAGAAATCGCTTGAACCCAGAGTGTGAGGTTGCAATGACTGA 2062
QY 1441 GATCGTGCCTTGCACCTCCAGCTGGGCAACAAGCAAGACTCTGTCTCAGAAAAAAA 1500
DB 2063 GATCGTGCCTTGCACCTCCAGCTGGGCAACAAGCAAGACTCTGTCTCAGAAAAAAA 2122
QY 1501 AAAAAAAGAGAGAGAGAGAAAGAAACATATTTTGGAGAGAGATGGGAGAGCAT 1560
DB 2123 AAAAAAAGAGAGAGAGAGAAAGAAACATATTTTGGAGAGAGATGGGAGAGCAT 2182
QY 1561 TGCAGAGAAATGTGCTTTATCCAAACAATGTAAAGAGCAATAAGGATCCCTATTTG 1620
DB 2183 TGCAGAGAAATGTGCTTTATCCAAACAATGTAAAGAGCAATAAGGATCCCTATTTG 2242
QY 1621 TCTCTTTGTGTCTATTTGTCTTAACAATGTCTTTGACAGTGAAGAAAAATTTCAAG 1680

Db 2243 TCTCTTTTGGTGTATTTGTCCTTAACTGTCCTTTGACAGTGAAGAAAAATATCAGA 2302
Qy 1681 ATAAACATATCCCTGTCCTGTTATTTACTAGCAACCTTGCAATGAAGAAGAGATCC 1740
Db 2303 ATAAACATATCCCTGTCCTGTTATTTACTAGCAACCTTGCAATGAAGAAGAGATCC 2362
Qy 1741 ACAGAAACCTGTAATGACAACTGCTATTTATCTTATGTAACATAGTTGTA 1800
Db 2363 ACAGAAACCTGTAATGACAACTGCTATTTATCTTATGTAACATAGTTGTA 2422
Qy 1801 AGAGTAAATATGTTACTGATGATTCATTTATTTATTTATTTATTTGCGTATG 1860
Db 2423 AGAGTAAATATGTTACTGATGATTCATTTATTTATTTATTTATTTGCGTATG 2482
Qy 1861 ATTTTATATTAACATGATTTCTTTCTGATATATGAAATGAGTCTCAAGCTTATA 1920
Db 2483 ATTTTATATTAACATGATTTCTTTCTGATATATGAAATGAGTCTCAAGCTTATA 2542
Qy 1921 AATTATATCTTTAGAAATGATTCATAAACAAGATGTAATTGTAACATGAGTAAT 1980
Db 2543 AATTATATCTTTAGAAATGATTCATAAACAAGATGTAATTGTAACATGAGTAAT 2602
Qy 1981 GGTGCTAGAGAGCCATTTCTCTGATTTTGTAGTAACTTTATGACAGCAATTTGCTC 2040
Db 2603 GGTGCTAGAGAGCCATTTCTCTGATTTTGTAGTAACTTTATGACAGCAATTTGCTC 2662
Qy 2041 TGGCTCATCTTCAATCAGTTAAATATGATGATTAATTTTGAAGCTGTGAAGTAAA 2100
Db 2663 TGGCTCATCTTCAATCAGTTAAATATGATGATTAATTTTGAAGCTGTGAAGTAAA 2722
Qy 2101 TACCAAAATTAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160
Db 2723 TACCAAAATTAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2782
Qy 2161 GGAATTAACCTG 2172
Db 2783 GGAATTAACCTG 2794

RESULT 3
US-09-966-880A-10
; Sequence 10, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FaSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 6564
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-966-880A-10

Query Match 100.0%; Score 2172; DB 9; Length 6564;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCCTGTATAGGTTGATATCTTACAGACGATTTGCTACTTTGGGACTTTGATAGCA 60
Db 3741 CCCCTGTATAGGTTGATATCTTACAGACGATTTGCTACTTTGGGACTTTGATAGCA 3800

Qy 61 CTTCCAGGATGTCAACACGATGAATATCTGCTGAGACAGTGGATTAACAGT 120
Db 3801 CTTCCAGGATGTCAACACGATGAATATCTGCTGAGACAGTGGATTAACAGT 3860
Qy 121 CCTTCAAGCTCTCTGTTTTTATCTTCAACTCTCACTTTCTTTAGATTTACAGAAA 180
Db 3861 CCTTCAAGCTCTCTGTTTTTATCTTCAACTCTCACTTTCTTTAGATTTACAGAAA 3920
Qy 181 AATTATATATGACTCTTTTAAAAAGATCTATGCTTGTAAATAGAGAGAACAGAGT 240
Db 3921 AATTATATATGACTCTTTTAAAAAGATCTATGCTTGTAAATAGAGAGAACAGAGT 3980
Qy 241 CTGGCAGGGAAGTCTGCAATTTGCTGAGTTTGAATGACAACTTGTCCCTACTGGA 300
Db 3981 CTGGCAGGGAAGTCTGCAATTTGCTGAGTTTGAATGACAACTTGTCCCTACTGGA 4040
Qy 301 ATAAAGAACTGACAGACCTGGAGCATCTTAAAGTCTCAAGTTTTCTATGACTTTA 360
Db 4041 ATAAAGAACTGACAGACCTGGAGCATCTTAAAGTCTCAAGTTTTCTATGACTTTA 4100
Qy 361 GGTAGATGAGACAGAAAGTATGATCTTAAAGATGCTGAGAGATCAATGTTTTTA 420
Db 4101 GGTAGATGAGACAGAAAGTATGATCTTAAAGATGCTGAGAGATCAATGTTTTTA 4160
Qy 421 TATCAACATCCTTTATTTATTTGATTCATTTGAGTTAACAGTGTGTAGTATGATTT 480
Db 4161 TATCAACATCCTTTATTTATTTGATTCATTTGAGTTAACAGTGTGTAGTATGATTT 4220
Qy 481 TCTATCTTTCCCTTGAAGTTTACTTTCAAGTAAACAAACTCTTCCATGAGGCATGA 540
Db 4221 TCTATCTTTCCCTTGAAGTTTACTTTCAAGTAAACAAACTCTTCCATGAGGCATGA 4280
Qy 541 TCTATAGACCTCTTAATAGAGATCTGAGTATGATGACCCCAACATCTTCCAAA 600
Db 4281 TCTATAGACCTCTTAATAGAGATCTGAGTATGATGACCCCAACATCTTCCAAA 4340
Qy 601 GCATTATATCCATATCATGCTGTATGTTTAAACAGAGAGATGTTTATGTTTG 660
Db 4341 GCATTATATCCATATCATGCTGTATGTTTAAACAGAGAGATGTTTATGTTTG 4400
Qy 661 TCAAAAGAAATGTTATGAGTGGGAGATGAGATATAGACCATGATGTCACCTTGA 720
Db 4401 TCAAAAGAAATGTTATGAGTGGGAGATGAGATATAGACCATGATGTCACCTTGA 4460
Qy 721 GCTACTTTATTAAGATCTTAAAAATGGGACAGAGACTGTGAACAAGCACCTTAATA 780
Db 4461 GCTACTTTATTAAGATCTTAAAAATGGGACAGAGACTGTGAACAAGCACCTTAATA 4520
Qy 781 TGGGTTGATGTCTGAGTGCAAATCTTTGGAACGCAACTCTTTTAAGAAATCCCT 840
Db 4521 TGGGTTGATGTCTGAGTGCAAATCTTTGGAACGCAACTCTTTTAAGAAATCCCT 4580
Qy 841 AATTGAAACACCCCAAACTTCAATATCATATTAATTAACAACATGGAAGAGTTG 900
Db 4581 AATTGAAACACCCCAAACTTCAATATCATATTAATTAACAACATGGAAGAGTTG 4640
Qy 901 CTGGAATGTGGGAGAGGAAATCTAATGCTCTGCTGAGGCTCTCTTCACTCAGAAATG 960
Db 4641 CTGGAATGTGGGAGAGGAAATCTAATGCTCTGCTGAGGCTCTCTTCACTCAGAAATG 4700
Qy 961 CCAATCAGGTCAGAGTTTCTCATTTTGTATGTGTATGATCTTCCAAAGATATAT 1020
Db 4701 CCAATCAGGTCAGAGTTTCTCATTTTGTATGTGTATGATCTTCCAAAGATATAT 4760
Qy 1021 TAACTATATTAAGAGATTTGTGCAAAACAGATGATTAAGCTGCGAACCGTGCACACG 1080
Db 4761 TAACTATATTAAGAGATTTGTGCAAAACAGATGATTAAGCTGCGAACCGTGCACACG 4820
Qy 1081 TCAATAGTTTAACTGCTGGAGAGTTGAGAGAGAGATGAGCTTGAACAAGGTGTTCA 1140
Db 4821 TCAATAGTTTAACTGCTGGAGAGTTGAGAGAGAGATGAGCTTGAACAAGGTGTTCA 4880

1141 GGCACGCTGGGCAATACAGATCCCTCTCTCAAAAAAAAAAAAAAAAAAGAA 1200
4881 GGCACGCTGGGCAATACAGATCCCTCTCTCAAAAAAAAAAAAAAAAAAGAA 4940
1201 GAGAGAGGCGCGCGCTGGCTCAAGCTGTAAATCCAGACATTTGGAGGCGGAGCC 1260
4941 GAGAGAGGCGCGCGCTGGCTCAAGCTGTAAATCCAGACATTTGGAGGCGGAGCC 5000
1261 GGGCGGATACCTGTGTGTGAGAGTTGAGACCAAGCTGGCCAAATGGCAAAACCCGT 1320
5001 GGGCGGATACCTGTGTGTGAGAGTTGAGACCAAGCTGGCCAAATGGCAAAACCCGT 5060
1321 CTGTACTCAAAATGCAAAATTTAGACGAGCGTGTAGAGAGGACCTGTAAATCCAGTAC 1380
5061 CTGTACTCAAAATGCAAAATTTAGACGAGCGTGTAGAGAGGACCTGTAAATCCAGTAC 5120
1381 TTGGAGGCTGAGGAGAGAAATCGCTTGAACCCAGAGGTGAGAGTTGCAAGTAACTGA 1440
5121 TTGGAGGCTGAGGAGAGAAATCGCTTGAACCCAGAGGTGAGAGTTGCAAGTAACTGA 5180
1441 GATCGTGGCTGTGACTCCAGCCTGGGCGAACAAGCAAGACTGTGTCTCAGAAAAAAA 1500
5181 GATCGTGGCTGTGACTCCAGCCTGGGCGAACAAGCAAGACTGTGTCTCAGAAAAAAA 5240
1501 AAAAAAAAAAGAGAGAGAGAGAGAAATATTTGGAGAGAGAGAGAGAGAGAGAGAGAT 1560
5241 AAAAAAAAAAGAGAGAGAGAGAGAAATATTTGGAGAGAGAGAGAGAGAGAGAGAGAT 5300
1561 TGCAAGAAATGTGCTTATCCAAACAAATGTAAGAGAGAGAGAGAGAGAGAGAGAT 1620
5301 TGCAAGAAATGTGCTTATCCAAACAAATGTAAGAGAGAGAGAGAGAGAGAGAGAT 5360
1621 TCTCTTGTGTGTATTTGTCCCTAACAATCTCTTTGAGAGAGAGAGAGAGAGAGAT 1680
5361 TCTCTTGTGTGTATTTGTCCCTAACAATCTCTTTGAGAGAGAGAGAGAGAGAGAT 5420
1681 ATACCAATATCCCTGTGCGGTATTAATCTAGCAACCTTGCATGAGAGAGAGAGAT 1740
5421 ATACCAATATCCCTGTGCGGTATTAATCTAGCAACCTTGCATGAGAGAGAGAGAT 5480
1741 ACGGAAATCTGAATGCAACCTGTCTAATTTTAATCTAATGAGAGAGAGAGAT 1800
5481 ACGGAAATCTGAATGCAACCTGTCTAATTTTAATCTAATGAGAGAGAGAGAT 5540
1801 AGAGTAAATATGTTACTTCAATGATTAATTTATTTATTTATTTATTTATTTATTTAT 1860
5541 AGAGTAAATATGTTACTTCAATGATTAATTTATTTATTTATTTATTTATTTATTTAT 5600
1861 ATTTTATTTAATCATGATTTCTTTCTGATAATTTGAATGAGAGAGAGAT 1920
5601 ATTTTATTTAATCATGATTTCTTTCTGATAATTTGAATGAGAGAGAGAT 5660
1921 AATTATTAATCTTGAATGATTTCTAATAACAAGTATATTTGTAATCTGAGAT 1980
5661 AATTATTAATCTTGAATGATTTCTAATAACAAGTATATTTGTAATCTGAGAT 5720
1981 GGTGTACGAGAGCATTTCTCTGATTTTATGAACTTTTGAAGAGAGAT 2040
5721 GGTGTACGAGAGCATTTCTCTGATTTTATGAACTTTTATGAGAGAGAT 5780
2041 TGCTCACTTTCAATCATGATTAATTAATGATTAATTTGAAAGCTGTGAAGAT 2100
5781 TGCTCACTTTCAATCATGATTAATTAATGATTAATTTGAAAGCTGTGAAGAT 5840
2101 TACCAATTAATTAATTAATGATTAATTAATGATTAATTTGAAAGCTGTGAAGAT 2160
5841 TACCAATTAATTAATTAATGATTAATTAATGATTAATTTGAAAGCTGTGAAGAT 5900
2161 GGAATTAATCTG 2172
5901 GGAATTAATCTG 5912

RESULT 4
US-09-966-880a-35
Sequence 35, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Taku
APPLICANT: Muramatsu, Masamichi
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966, 880A
PRIORITY FILING DATE: 2001-09-28
PRIORITY FILING DATE: 2000-03-28
PRIORITY FILING DATE: 1999-12-27
PRIORITY FILING DATE: 1999-06-24
PRIORITY FILING DATE: 1999-03-29
NUMBER OF SEQ. ID NOS: 36
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 35
LENGTH: 11204
TYPE: DNA
ORGANISM: Homo sapiens
US-09-966-880a-35
Query Match 100.0%; Score 2172; DB 9; Length 11204;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CCCTGTATGAGGTGATGATCTTACGAGAGCAATTTGATCTTGGAGCTTGTATGCA 60
8956 CCCGTATGAGGTGATGATCTTACGAGAGCAATTTGATCTTGGAGCTTGTATGCA 9015
61 CTTCAGAGATGTCACACAGATGAATATCTGCTGAAGACAGAGATTAATAACAGT 120
9016 CTTCAGAGATGTCACACAGATGAATATCTGCTGAAGACAGAGATTAATAACAGT 9075
121 CCTCAAGCTCTGCTGTTTATTTATCTTCACTCTCACTTTCTTAAGTTTACAGAAAA 180
9076 CCTCAAGCTCTGCTGTTTATTTATCTTCACTCTCACTTTCTTAAGTTTACAGAAAA 9135
181 ATATTTATATGACATCTTTTAAAGATCTATGCTTTGAATATGAGAGAGACACAGT 240
9136 ATATTTATATGACATCTTTTAAAGATCTATGCTTTGAATATGAGAGAGACACAGT 9195
241 CTGCGCAGAGAGCTGTCGCAATTTGTCAGATTTTGAATGACATTTGCCCTACAGGA 300
9196 CTGCGCAGAGAGCTGTCGCAATTTGTCAGATTTTGAATGACATTTGCCCTACAGGA 9255
301 ATACAGACCTGACAGAGCTGAGAGCATCTTAAGTGTCAAGCTTTTCTATGACTTTTA 360
9256 ATACAGACCTGACAGAGCTGAGAGCATCTTAAGTGTCAAGCTTTTCTATGACTTTTA 9315
361 GGTAGATGAGAGAGAGATGATCTTAAAGATGAGAGATGAGAGATCAATGTTTTTA 420
9316 GGTAGATGAGAGAGAGATGATCTTAAAGATGAGAGATGAGAGATCAATGTTTTTA 9375
421 TATCAACATCTTTATTTATTTATGATCTTTGATTAAGTGTGTTAGATGATTTT 480
9376 TATCAACATCTTTATTTATTTATGATCTTTGATTAAGTGTGTTAGATGATTTT 9435
481 TCTATCTTTTCCCTGAGGTTTACTTCAAGTAAACAAACCTTCCATCAGGCGATGA 540
9436 TCTATCTTTTCCCTGAGGTTTACTTCAAGTAAACAAACCTTCCATCAGGCGATGA 9495
541 TCTATGAGAGCTCTATGAGAGATCTGGGATGATGAGAGAGAGAGAGAGAGAGAGAT 600
9496 TCTATGAGAGCTCTATGAGAGATCTGGGATGATGAGAGAGAGAGAGAGAGAGAGAT 9555
601 GCATTAATATCAATCAATGCGCTGATGTTTATATGAGAGAGAGAGATTTTATGTTTG 660

Db 9556 GCATTATATCCAAATCATGCGCTGTATGTTTAAATCAGCAGAGACATGTTTATGTTTG 9615
 Qy 661 TACAAAAGAAATTTGTTATGGTGGGATGAGAGATAGACCATGAGTCACTTCAA 720
 Db 9616 TACAAAAGAAATTTGTTATGGTGGGATGAGAGATAGACCATGAGTCACTTCAA 9675
 Qy 721 GCTACTTTAAATTAAGATCTTAAATGAGGAGAGAGTGTAAACAGACACCTTAATA 780
 Db 9676 GCTACTTTAAATTAAGATCTTAAATGAGGAGAGAGTGTAAACAGACACCTTAATA 9735
 Qy 781 TGGGTGATGTCGTGAAGTGAAGTCTTCTGAAAACGCAAACTCTTTTAAAGAAATGCTT 840
 Db 9736 TGGGTGATGTCGTGAAGTGAAGTCTTCTGAAAACGCAAACTCTTTTAAAGAAATGCTT 9795
 Qy 841 AATTTAGAAACCCCAAACTTCAATCTCAATCTTAATTTAGCAACAAATGGAAGAAATG 900
 Db 9796 AATTTAGAAACCCCAAACTTCAATCTCAATCTTAATTTAGCAACAAATGGAAGAAATG 9855
 Qy 901 CTGGAATGTTGGGAGAGAGAAATGATTTGCTCGTGGGTCTCTCATCTCAGAAATG 960
 Db 9856 CTGGAATGTTGGGAGAGAGAAATGATTTGCTCGTGGGTCTCTCATCTCAGAAATG 9915
 Qy 961 CCAATCAGTCAAGGTTTGTCTACATTTTGTATGTGTGTATGCTTCTCCAAAGGTATAT 1020
 Db 9916 CCAATCAGTCAAGGTTTGTCTACATTTTGTATGTGTGTATGCTTCTCCAAAGGTATAT 9975
 Qy 1021 TAACATATTAAGAGAGTGTGAACAAACAGAAATGATTAAGCTGCAACCGTGGCAGACC 1080
 Db 9976 TAACATATTAAGAGAGTGTGAACAAACAGAAATGATTAAGCTGCAACCGTGGCAGACC 10035
 Qy 1081 TCATAGTTCTAGTGTGTTGGAGAGTGTGAAGAGAGAGATGCTTGAACACGAGTTCAA 1140
 Db 10036 TCATAGTTCTAGTGTGTTGGAGAGTGTGAAGAGAGAGATGCTTGAACACGAGTTCAA 10095
 Qy 1141 GGCACAGCTGGGCAACATTAACAAGATCTGTCTCTCAAAAAAATTTTAAAAAAGAA 1200
 Db 10096 GGCACAGCTGGGCAACATTAACAAGATCTGTCTCTCAAAAAAATTTTAAAAAAGAA 10155
 Qy 1201 GAGAGAGAGCGGCGCTGTGTGCTCAAGCTGTATCCAGACCTTTGGAGGCGGAGCC 1260
 Db 10156 GAGAGAGAGCGGCGCTGTGTGCTCAAGCTGTATCCAGACCTTTGGAGGCGGAGCC 10215
 Qy 1261 GGGCGGATCACTGTGTGTGAAGAGTGTGAACAGAGCTGTGCCAATGGCAAAACCCGT 1320
 Db 10216 GGGCGGATCACTGTGTGTGAAGAGTGTGAACAGAGCTGTGCCAATGGCAAAACCCGT 10275
 Qy 1321 CTGTACTCAAAATGCAAAATTTAGCAAGGCGTGTGAGGAGACCTGTATCCAGACTAC 1380
 Db 10276 CTGTACTCAAAATGCAAAATTTAGCAAGGCGTGTGTGAGGAGACCTGTATCCAGACTAC 10335
 Qy 1381 TTGGAGAGCTGAGGAGAGATCGCTTGAACCCAGAGAGTGAAGTTGCAATTAAGCTGA 1440
 Db 10336 TTGGAGAGCTGAGGAGAGATCGCTTGAACCCAGAGAGTGAAGTTGCAATTAAGCTGA 10395
 Qy 1441 GATGTGCGCTGTGCACTCCAGCTGTGGCGCAAGAGCAAGATCTGTCTCAAAAAA 1500
 Db 10396 GATGTGCGCTGTGCACTCCAGCTGTGGCGCAAGAGCAAGATCTGTCTCAAAAAA 10455
 Qy 1501 AAAAAAGAT 1560
 Db 10456 AAAAAAGAT 10515
 Qy 1561 TGAAGAGAAATGAGTGTATCCAAACAAATGTAAGAGCAGATTAAGGATCCCTATTTG 1620
 Db 10516 TGAAGAGAAATGAGTGTATCCAAACAAATGTAAGAGCAGATTAAGGATCCCTATTTG 10575
 Qy 1621 TCTCTTTGGTGTCTATTTGCTCTTAACAACCTGTCTTTGACAGTGAAGAAATATTCAGA 1680
 Db 10576 TCTCTTTGGTGTCTATTTGCTCTTAACAACCTGTCTTTGACAGTGAAGAAATATTCAGA 10635
 Qy 1681 AATAACATATCCCTGTGCGCTTATTAACCTAGCAACCTTGCATTAAGAGTGAAGATCC 1740
 Db 10636 AATAACATATCCCTGTGCGCTTATTAACCTAGCAACCTTGCATTAAGAGTGAAGATCC 10695

Qy 1741 ACAGAAAACTTGAATGCAACAATGCTATTTTAACTTATTTGATCAATAGTTGTAAA 1800
 Db 10696 ACAGAAAACTTGAATGCAACAATGCTATTTTAACTTATTTGATCAATAGTTGTAAA 10755
 Qy 1801 AGAGTAAATAATTTGTTACTTCAATGATATCATTTATTTTATTTATTTTGGCTAATG 1860
 Db 10756 AGAGTAAATAATTTGTTACTTCAATGATATCATTTATTTTATTTATTTTGGCTAATG 10815
 Qy 1861 AATTTTATTAACATGATTCCTTTCTGATATATTTGAATGAAATGAGTCTCAAGTTTCA 1920
 Db 10816 AATTTTATTAACATGATTCCTTTCTGATATATTTGAATGAAATGAGTCTCAAGTTTCA 10875
 Qy 1921 AATTTTATTAACATGATTCCTTTCTGATATATTTGAATGAAATGAGTCTCAAGTTTCA 1980
 Db 10876 AATTTTATTAACATGATTCCTTTCTGATATATTTGAATGAAATGAGTCTCAAGTTTCA 10935
 Qy 1981 GGTGTACGAGAGCATTTCTCTGATTTTGTAACTTTTATGACAGAAATTTGCTTC 2040
 Db 10936 GGTGTACGAGAGCATTTCTCTGATTTTGTAACTTTTATGACAGAAATTTGCTTC 10995
 Qy 2041 TGGCTCACTTCAATCAGTAAATTAATGATTAATTTTGAAGCTGTGAAGTAA 2100
 Db 10996 TGGCTCACTTCAATCAGTAAATTAATGATTAATTTTGAAGCTGTGAAGTAA 11055
 Qy 2101 TACCAATTAATAATTAATAAGTATTTATATGAAGTTAAATTAATAATCAGTATGAT 2160
 Db 11056 TACCAATTAATAATTAATAAGTATTTATATGAAGTTAAATTAATAATCAGTATGAT 11115
 Qy 2161 GGAATTAACCTG 2172
 Db 11116 GGAATTAACCTG 11127

RESULT 5
 US-09-764-891-5477
 ; Sequence 5477, Application US/09764891
 ; Publication No. US20030077808A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC006
 ; CURRENT APPLICATION NUMBER: US/09/764,891
 ; Prior application data removed - consult PAM or file wrapper
 ; NUMBER OF SEQ ID NOS: 10231
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5477
 ; LENGTH: 13862
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-764-891-5477

Query Match 13.6%; Score 295.2; DB 10; Length 13862;
 Best Local Similarity 74.6%; Pred. No. 3.5e-5;
 Matches 397; Conservative 0; Mismatches 133; Indels 2; Gaps 2;

Qy 1045 AACAGATGATTAAGTGTGGAACCGTGGCAGACGCTCATATGTTCTAGCTGTTGGAGG 1104
 Db 1508 AAATTAATAAGTGTGCGGCTGTGTGCTCAATGCTGTATCCAGCACTTTGGAGG 1567
 Qy 1105 TTGAGAGAGAGAGTGGCTTGAACACAGTGTTCAGAGCCAGCTGGGCAACATPAACAG 1164
 Db 1568 CTGAGGAGAGAGATCACTTGAAGTCAAGAGTTGAACACAGCTGTGCCAATGCGGAA 1627
 Qy 1165 ATCTGTCTCCAAAAAATTTTAAAAAAGAAAGAGAGAGGCGGCGTGTGCGC 1223
 Db 1628 ACCCATCTCTAATAATAATTAATAAGTATGCGGAGATGAGCGCGGAGATGAGGCGC 1687
 Qy 1224 TCAGCGCTGATCCAGCACTTTGGAGGCGGAGCCGAGCGAGATCACTGTGTGTCAGA 1283
 Db 1688 TCATGCGCTGTATCCAGCACTTTGGAGGCGGAGCGTGTGCGGATCACTGTGAGTCA 1747

QY 1284 GTTGAAGACAGCTGGCCACAATGGCAAAACCCGCTGTACTCAAAATGCAAAATTA 1343
|||
Db 1748 GTTCAAGACAGAGCTGGCAACAATGGCAAAACCCGCTGTACTCAAAATTA 1807
|||
QY 1344 GCCAGGCGTGTGACAGGCACTGTATCCAGCTACTTGGAGGCTGAGGAGAGAAAT 1403
|||
Db 1808 GCCAGGCTTATGTGTGGGCGCTTATATCCAGCTACTTGGAGGCTGAGGAGAGAAAT 1867
|||
QY 1404 CGCTTAACCCAGAGAGTGGAGGTGACATTAAGCTGAGATCGTGGCTGGCACTCAGGC 1463
|||
Db 1868 CGTTTAAGCCAGAGAGCGAGAGTTTGACCCAGCCAGATGTGGCCACTGCACTCCAGCC 1927
|||
QY 1464 TGGGCGACAAAGACCAAGACTCTGTCTCAGAAAAAAGAGAGAGAGAGAGAA 1523
|||
Db 1928 TGGGCGCATAA-AGCGAGACTCTGTCTCAGAAAAAAGAGAGAGAGAGAGAA 1986
|||
QY 1524 AGAGAAACAATTTTGGGAGAGAAAGATGGGGAAGCAATTGCAAGAAATTTGTG 1575
|||
Db 1987 AGAAAAAATTAGCCAGAGCTGTGTGGCAATGCACTGTGTCTAGCTACTTGGG 2038
|||

RESULT 6

US-09-764-891-10204
; Sequence 10204, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10204
; LENGTH: 13862
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-10204

Query Match 13.6%; Score 295.2; DB 10; Length 13862;
Best Local Similarity 74.6%; Pred. No. 3.5e-51;
Matches 397; Conservative 0; Mismatches 133; Indels 2; Gaps 2;

QY 1045 AAACAGATGATTAAGCTGGCAACCGTGGCAACCGTCAATGTTAGCTGTGGAGG 1104
|||
Db 1508 AAATTAATTAAGCTGGCTGGCTGTGTGTGCTATGCTGTATCCAGACATTTGGAGG 1567
|||
QY 1105 TTGAGAGGAGAGATGGCTTTGAACACAGGTGTTCAAGGCCAGCTGGGCAATTAAG 1164
|||
Db 1568 CTGAGGCAAGAGATCACTTGAGTCAAGAGTTTGACCAACCTGGCCAAATGAGGAA 1627
|||
QY 1165 ATCTGTCTCTCAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1223
|||
Db 1628 ACCCATCTCTCTAATAAAATACAAAGTTAGCCAGGCAATGGCCGAGTGTGGTGGC 1687
|||
QY 1224 TCAGGCTGTATATCCAGCACTTTGGAGGCGAGCGGCGGATCACTGTGTGAGGA 1283
|||
Db 1688 TCATGCTGTATATCCAGCACTTTGGAGGCGGAGTGGCGGATCACTGTGTGAGGA 1747
|||
QY 1284 GTTGAAGACAGCTGGCCACAATGGCAAAACCCGCTGTACTCAAAATGCAAAATTA 1343
|||
Db 1748 GTTCAAGACAGAGCTGGCAACAATGGCAAAACCCGCTGTACTCAAAATTA 1807
|||
QY 1344 GCCAGGCGTGTGACAGGCACTGTATCCAGCTACTTGGAGGCTGAGGAGAGAAAT 1403
|||
Db 1808 GCCAGGCTTATGTGTGGGCGCTTATATCCAGCTACTTGGAGGCTGAGGAGAGAAAT 1867
|||
QY 1404 CGTTTAAGCCAGAGAGTGGAGGTGACATTAAGCTGAGATCGTGGCTGGCACTCAGGC 1463
|||
Db 1868 CGTTTAAGCCAGAGAGCGAGAGTTTGACCCAGCCAGATGTGGCCACTGCACTCCAGCC 1927
|||
QY 1464 TGGGCGACAAAGACCAAGACTCTGTCTCAGAAAAAAGAGAGAGAGAGAGAA 1523
|||

Db 1928 TGGGCGATTA-AGCGAGACTCTGTCTCAGAAAAAAGAGAGAGAGAGAGAA 1986
|||
QY 1524 AGAGAAACAATTTTGGGAGAGAAAGATGGGGAAGCAATTGCAAGAAATTTGTG 1575
|||
Db 1987 AGAAAAAATTAGCCAGAGCTGTGTGGCAATGCACTGTGTCTAGCTACTTGGG 2038
|||

RESULT 7

US-10-205-428-1003
; Sequence 1003, Application US/10205428
; Publication No. US20030108907A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA117C1
; CURRENT APPLICATION NUMBER: US/10/205,428
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 09/764,892
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1019
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1003
; LENGTH: 13862
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-428-1003

Query Match 13.6%; Score 295.2; DB 14; Length 13862;
Best Local Similarity 74.6%; Pred. No. 3.5e-51;
Matches 397; Conservative 0; Mismatches 133; Indels 2; Gaps 2;

QY 1045 AAACAGATGATTAAGCTGGCAACCGTGGCAACCGTCAATGTTAGCTGTGGAGG 1104
|||
Db 1508 AAATTAATTAAGCTGGCTGGCTGTGTGTGCTATGCTGTATCCAGACATTTGGAGG 1567
|||
QY 1105 TTGAGAGGAGAGATGGCTTTGAACACAGGTGTTCAAGGCCAGCTGGGCAATTAAG 1164
|||
Db 1568 CTGAGGCAAGAGATCACTTGAGTCAAGAGTTTGACCAACCTGGCCAAATGAGGAA 1627
|||
QY 1165 ATCTGTCTCTCAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1223
|||
Db 1628 ACCCATCTCTCTAATAAAATACAAAGTTAGCCAGGCAATGGCCGAGTGTGGTGGC 1687
|||
QY 1224 TCAGGCTGTATATCCAGCACTTTGGAGGCGAGCGGCGGATCACTGTGTGAGGA 1283
|||
Db 1688 TCATGCTGTATATCCAGCACTTTGGAGGCGAGGTTGGCGGATCACTGTGTGAGGA 1747
|||
QY 1284 GTTGAAGACAGCTGGCCACAATGGCAAAACCCGCTGTACTCAAAATGCAAAATTA 1343
|||
Db 1748 GTTCAAGACAGAGCTGGCAACAATGGCAAAACCCGCTGTACTCAAAATTA 1807
|||
QY 1344 GCCAGGCGTGTGACAGGCACTGTATATCCAGCTACTTGGAGGCTGAGGAGAGAAAT 1403
|||

Db 1808 GCCAGCTTTATGTTAGGCGCTATATATCCAGACTCTGGAAGGCTGAGGACAGAGAT 1867
QY 1404 CGCTTGAACCCAGAGAGTGGAGGTTGCAGTAACTGAGATGTGCGCTTGACTCCAGCC 1463
Db 1868 CGTTTGAAGCCAGAGAGGAGTTTGACCCGAGCCGAGATTTGTGCACTGCATCCAGCC 1927
QY 1464 TGGGCGACAGAGCAAGACTGTCTCAGAAAAAAGAGAGAGAGAGAGAA 1523
Db 1928 TGGGCGATTA-AGCAGAGACTGTCTCAGAAAAAAGAGAGAGAGAGAGAA 1986
QY 1524 AGAGACATATTTTGGAGAGAGAGATGGGAGAGATTTGCAAGAAATTTGTG 1575
Db 1987 AGAAAAAATTAGCCAGCGCTGTGGCATGACCTGTAACTTAACTTAACTTGGG 2038

RESULT 8

US-09-764-891-6094
; Sequence 6094, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6094
; LENGTH: 7809
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-6094

Query Match 13.5%; Score 292.6; DB 10; Length 7809;
Best Local Similarity 77.2%; Pred. No. 9.3e-51;
Matches 385; Conservative 0; Mismatches 104; Indels 10; Gaps 2;

QY 1044 AAAACAGATGATTAAGTGGCGAACCGTGCACACGCTCATATGTTCTAGCTCTTGGAG 1103
Db 1081 AAAAAGTTAAATCAAGCTGTGGCATGTGTGCTCAGCCCTAATATCCAGCACTTTGGAG 1140
QY 1104 GTTGAAGAGGAGATGGCTTGAACAAGGTGTTCAAGGCCAGCTGGCAACATPACAA 1163
Db 1141 GCTGAAGGAGGAGATCACTTGAAGCCAGAGCTCAAGACCAACCTGCACAAGATGATGA 1200
QY 1164 GATCTGTCTCTCAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1223
Db 1201 AACCTGTCTCTCAATTTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1251
QY 1224 TCACGCTGTAAATCCAGCACTTTGGAGGCGAGCGGAGGATCACTGTGTGAGGA 1283
Db 1252 TCACGCTGTAAATCCAGCACTTTGGAGGCGAGCGGAGGATCACTGTGTGAGGA 1311
QY 1284 GTTGAAGAGGAGCTGGCAACATGGCAAAACCCGCTGTACTCAAAAATGCAAAATTA 1343
Db 1312 GTTCAAGAGAGAGCTGACCAACATGTGAAACCCCACTGTACAAAATTAATAAATA 1371
QY 1344 GCCAGGCGTGTAGCAGGCACTGTAAATCCAGCTACTTGGAGGCTGAGGCGAGAGAT 1403
Db 1372 GCTGGGCGTGTGCTCATGCTGTAAATCCAGCTACTTGGAGGCTGAGGCGAGAGAT 1431
QY 1404 CGCTTGAACCCAGAGAGTGGAGGTTGAGTAACTGAGATCGAGCGCTGCATCCAGCC 1463
Db 1432 CGATTGAACCTAGAGGCAAGGTTGATGAGTGAATCGCACTTGCATCCAGAC 1491
QY 1464 TGGGCGACAGAGG-AGAAGCTGTCTCAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAG 1522
Db 1492 TGGGCGACAGAGACAAAACTCACTCTCAAAAAAAGAGAGAGAGAGAGAGAGAGAGAG 1551
QY 1523 AAGAGAACATATTTTGGGA 1541
Db 1552 GTTAAACAAAGTATTTGCGA 1570

RESULT 9

US-09-764-891-7364
; Sequence 7364, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7364
; LENGTH: 32249
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7364

Query Match 13.1%; Score 284.4; DB 10; Length 32249;
Best Local Similarity 75.7%; Pred. No. 9.9e-49;
Matches 365; Conservative 0; Mismatches 116; Indels 1; Gaps 1;

QY 1060 GCTGCGAACCGTGGCAACGCTCATATGTTCTAGCTCTTGGAGGTTGAGAGAGAGAT 1119
Db 15040 GCTGTGACAGTGTGCTCATGCTCTGTAACTCAAGCACTTGGAGATCTAGGCGGTGAGT 15099
QY 1120 GCTTGAACAGAGTGTCAAGGCCAGCTGGCAACATPACAAAGTCTGTCTCAAA 1179
Db 15100 CGCTTGAAGCCAGAGAGTCAAGACCAAGCTGGCAACATPACAGCCCATCTPAAA 15159
QY 1180 AAAAAAAG 1239
Db 15160 AAATTAATAATTAATAAAGCTTTAGGCGCAGGCGGTGTGCTCAGCCTGTAAATCC 15219
QY 1240 AGCACTTTGGAGAGCGGCGGCGGAGATCACTGTGTGAGAGTTGAGACCAAGCTG 1299
Db 15220 AGCCCTTTGGAGAGCGGCGGCGGAGATCACTGTGTGAGAGTTGAGACCAAGCTG 15279
QY 1300 GCCAAGATGGCAAAACCCGCTGTACTCAAAATGCAAAATTTAGCAGGCGGTGAGCA 1359
Db 15280 GCCAAGATGGCAAAATCTCACTCTACTGAAATTAACCAAGCATGTGAGGCG 15339
QY 1360 GGCACCTGTAAATCCAGCTACTTGGAGGCTGAGAGAGAGATCGCTTGAACCCAGAG 1419
Db 15340 GCGCCTGTAAATCCAGCTACTTGAAGGATGAGGAGAGAGATCGTGAATCCAGAG 15399
QY 1420 GTGAGGTTGAGTAACTGAGATCGTCCGCTTGCATCCAGGCTGGGCGCAAGAGCA 1479
Db 15400 GAGGACATGTGAGAGCCAGAGTGTGCGCACTGCACTCAGGCTGGTAC-AGAGCCA 15458
QY 1480 GACTGTGTCTCAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1539
Db 15459 GAGTCAATCTCAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15518
QY 1540 GA 1541
Db 15519 TA 15520

RESULT 10

US-09-962-436-292
; Sequence 292, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sigma
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; CURRENT FILING DATE: 2001-09-25

PRIOR APPLICATION NUMBER: US/60/235,082
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/234,924
PRIOR FILING DATE: 2000-09-25
NUMBER OF SEQ ID NOS: 568
SOFTWARE: Patent in version 3.0
SEQ ID NO 292
LENGTH: 65608
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: n=a,t,g or c
US-09-962-436-292

Query Match 13.0%; Score 281.4; DB 9; Length 65608;
Best Local Similarity 81.2%; Pred. No. 6e-48;
Matches 376; Conservative 0; Mismatches 81; Indels 6; Gaps 4;

QY 1060 GCTGCAACCGTGGCAACGCTCATAGTTCTAGCTGCTTGGAGGTTGAGAGGAGAT 1119
DB 43089 GCCGGGTGTGTGGCTCAAGCTGTATCCAGCACTTGGAGTGTGAGGAGAGAT 43148
QY 1120 GGCTTGAACAGAGTGTGAGGCGAGCGCTGGCAACATTAACAAGATCTGTCTCAAA 1179
DB 43149 CCCTTGAAGCCAGAAATTCAA-GCCAGCTGGGCAATGTAGGAGAGCGCATCTCTACTA 43207
QY 1180 AAAAAAAAAAAAAAAAAAAGA--AAGAGAGAGGCGCGGCGTGTGCTCACGCTGTATC 1237
DB 43208 AAAAAAAAAAAAAAAAAAAGAGAGAAATTAGCCCGGGGTGTGTGCTCAAGCTGTATC 43267
QY 1238 CCAGCACTTTGGAGGCGGAGCGGCGGATCACTGTGTGTCAGAGATTGAGACCAACC 1297
DB 43268 CCAGCACTTTGGAGGCGTGAAGGAGGAGCGGATCAC--GAGGTCAAGATTGAGACCAACC 43325
QY 1298 TGGCCAACTGGCAAAACCCGCTGTGTACTCAAAATGCAAAATTAAGCCAGCGGTGTAG 1357
DB 43326 TGGCCAAAGATGTGAAGAAACCTGTCTCTACTTAAATAATTAAGCCAGCGGTGTAG 43385
QY 1358 GAGGACCTGTATCCAGCTACTTGGAGGCTGAGGAGAGAAATCGTTGAACCCAGG 1417
DB 43386 CGGCGCTGTATCCAGCTACTTCAAGAGGCTTGAAGGAGAGAAATTCCTTCAACCTGGG 43445
QY 1418 AGGTGAGGTTGCAATGAGATCGTCCGTTGCACTCCAGGCTGGCGCAAGAGAC 1477
DB 43446 AGGCGAGGTTGCAAGTGAAGCAAGATCGTCCAGCTGTATCCAGGCTGGGTGACAA-AGC 43504
QY 1478 AAGACTCTGTCTCAGAAAAAAAAAAAAAAAAAGAGAGAGAGA 1520
DB 43505 AAGACTCCATCTCAAAAAAAAAAAAAAAAAAGAGAGAGA 43547

RESULT 11
US-09-962-832-119
Sequence 119, Application US/09962832
Patent No. US2002010821A1
GENERAL INFORMATION:
APPLICANT: Emnet, Reinhard
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
FILE REFERENCE: 689290-74
CURRENT APPLICATION NUMBER: US/09/962,832
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/60/235,077
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,280
PRIOR FILING DATE: 2000-09-25
NUMBER OF SEQ ID NOS: 259
SOFTWARE: Patent in version 3.0
SEQ ID NO 119
LENGTH: 65608
TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: n=a,t,g or c
US-09-962-832-119

Query Match 13.0%; Score 281.4; DB 9; Length 65608;
Best Local Similarity 81.2%; Pred. No. 6e-48;
Matches 376; Conservative 0; Mismatches 81; Indels 6; Gaps 4;

QY 1060 GCTGCAACCGTGGCAACGCTCATAGTTCTAGCTGCTTGGAGGTTGAGAGGAGAT 1119
DB 43089 GCCGGGTGTGTGGCTCAAGCTGTATCCAGCACTTGGAGTGTGAGGAGAGAT 43148
QY 1120 GGCTTGAACAGAGTGTGAGGCGAGCGCTGGCAACATTAACAAGATCTGTCTCAAA 1179
DB 43149 CCCTTGAAGCCAGAAATTCAA-GCCAGCTGGGCAATGTAGGAGAGCGCATCTCTACTA 43207
QY 1180 AAAAAAAAAAAAAAAAAAAGA--AAGAGAGAGGCGCGGCGTGTGCTCACGCTGTATC 1237
DB 43208 AAAAAAAAAAAAAAAAAAAGAGAGAAATTAGCCCGGGGTGTGTGCTCAAGCTGTATC 43267
QY 1238 CCAGCACTTTGGAGGCGGAGCGGCGGATCACTGTGTGTCAGAGATTGAGACCAACC 1297
DB 43268 CCAGCACTTTGGAGGCGTGAAGGAGGAGCGGATCAC--GAGGTCAAGATTGAGACCAACC 43325
QY 1298 TGGCCAACTGGCAAAACCCGCTGTGTACTCAAAATGCAAAATTAAGCCAGCGGTGTAG 1357
DB 43326 TGGCCAAAGATGTGAAGAAACCTGTCTCTACTTAAATAATTAAGCCAGCGGTGTAG 43385
QY 1358 GAGGACCTGTATCCAGCTACTTGGAGGCTGAGGAGAGAAATCGTTGAACCCAGG 1417
DB 43386 CGGCGCTGTATCCAGCTACTTCAAGAGGCTTGAAGGAGAGAAATTCCTTCAACCTGGG 43445
QY 1418 AGGTGAGGTTGCAATGAGATCGTCCGTTGCACTCCAGGCTGGCGCAAGAGAC 1477
DB 43446 AGGCGAGGTTGCAAGTGAAGCAAGATCGTCCAGCTGTATCCAGGCTGGGTGACAA-AGC 43504
QY 1478 AAGACTCTGTCTCAGAAAAAAAAAAAAAAAAAGAGAGAGAGA 1520
DB 43505 AAGACTCCATCTCAAAAAAAAAAAAAAAAAAGAGAGAGA 43547

RESULT 12
US-09-964-531-180
Sequence 180, Application US/09954531
Patent No. US20020165180A1
GENERAL INFORMATION:
APPLICANT: Weaver, Zoe
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Ca
FILE REFERENCE: 689290-77
CURRENT APPLICATION NUMBER: US/09/954,531
PRIOR FILING DATE: 2002-05-02
PRIOR APPLICATION NUMBER: US/60/233,133
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,009
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,034
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,509
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: US/60/234,567
PRIOR FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 1392
SOFTWARE: Patent in version 3.0
SEQ ID NO 180
LENGTH: 65608
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(65608)
OTHER INFORMATION: n=a,t,g or c

US-09-954-531-180

| | | | | |
|-----------------------|--------------|------------------|----------------|---------------|
| Query Match | 13.0%; | Score 281.4; | DB 9; | Length 65608; |
| Best Local Similarity | 81.2%; | Pred. No. 6e-48; | | |
| Matches 376; | Conservative | 0; | Mismatches 81; | Indels 6; |
| | | | | Gaps 4; |

| | | | | | |
|----|-------|--|---|------------------------|------|
| QY | 1060 | GCCTGCMAACCGTGGCACAACGCTCA | TAATCTAGCTGTGGGAGGTTGAGAGGAGAT | 1119 | |
| Db | 43089 | GCCGGGTTGTGGTCTACGCTCTGTAAT | CCACGACCTTTGGAAGCTGTAGGACGAGGAT | 43148 | |
| QY | 1120 | GGCTTGAACA | CAGGTGTTCAGGCGACGCTGGGCAAT | TAACAAGATCCTGTCTCTCAAA | 1179 |
| Db | 43149 | CCCTTGAAGCCAGAAAGTTCAA | -GCCACCTGGGCAATGTGAGGAGACGCACATCTCTCACTA | 43207 | |
| QY | 1180 | AAAAAAAAAAAAAAAAA | GA--MAGAGAGAGGCGCGGCGTGTGGCTACAGCGCTGTAAATC | 1237 | |
| Db | 43208 | AAAAAAAAAAAAAAAAA | AGAGAAATTTAGCCGGGCGTGTGGGCTCAAGCGCTGTAAATC | 43267 | |
| QY | 1238 | CCAGCACTTTGGAGGCGCCGAGCCGGGCGGATCA | CCTGTGCTCAGAGATTGAGACCAAGCC | 1297 | |
| Db | 43268 | CCAGCACTTTGGAGGCGTGAAGGCAAGGCGATCAAC | -GAGTTCAGAGATTGAGACCAAGCC | 43325 | |
| QY | 1298 | TGGCCCAATGCGCAAAACCCCGCTTGTATCTCAAAAT | TCCAAAAATTTAGCCAGGCGTGTAG | 1357 | |
| Db | 43326 | TGGCCCAATGCTGAACCCCGCTTCTTCACTAAAAAT | TCCAAAAATTTAGCCAGGCGCGTGTAG | 43385 | |
| QY | 1358 | CAGGCACTGTAAATCCGACGACTTGGAGGCTGAGCGAAGAAAT | CGCTTGAACCCAGG | 1417 | |
| Db | 43386 | CGGGCGCTGTAAATCCAGGCTAATCAGAGGCTGAGCGAAGAAAT | TTCCTTCAACCTGGG | 43445 | |
| QY | 1418 | AGGTGAGGTTGCAAGTAAGCTGAAGATGTGCGCTTGCACTTC | CAGGCTTGGGCGCAAGAGCC | 1477 | |
| Db | 43446 | AGGCGGAAGTTGCAAGTGAAGCCAAATGCTGTCACTGTACCTCC | AGCCGCGGTGACAA-AGC | 43504 | |
| QY | 1478 | AAGACTGTGCTCAGAAAAA | AAAAAAAAAAGAGAGAGAGA | 1520 | |
| Db | 43505 | AAGACTTCATCTCAAAAAA | AAAAAAAAAAGAAAGAAAGAA | 43547 | |

```

RESULT 13
US-09-764-877-2857
; Sequence 2857, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; SEQ. ID NO 2857
; LENGTH: 11821
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2857

```

| | | | | |
|---------------------------|-------|-------------------|-----------|--------------|
| Query Match | 12.9% | Score 280.2 | DB 9 | length 11821 |
| Best Local Similarity | 74.4% | Pred. No. 4.5e-48 | | |
| Matches 386; Conservative | 0 | Mismatches 118 | Indels 15 | Gaps 27 |

| | | | |
|----|------|---|------|
| Qy | 1013 | AGGTATATTAACTATATTAAGAGGTTGTACAAAACAGATGATTAAGTCGAACCGG | 1072 |
| Db | 4599 | AGAAAATCGCAACTATGTTGGACGACAAAGAAATATTAATGAGTCGAGCCATGATG | 4657 |
| Qy | 1073 | GCAACAGCCTATAGTTCTAGCTGCTGGAGGTTGAGAGGAGATGCGCTTGAACACAG | 1133 |
| Db | 4659 | GCTCATGCCGGTATATCCACGACTTTAAGG-ATTTAGCGAGAGATCACTTGAGGCCAG | 4712 |
| Qy | 1133 | GTTGTCAGGCCACAGCTGGSCAACATACACAGATCTCTGTCCTCAAAAAAAAAAAAAA | 1193 |
| Db | 4718 | GAGTTCAAGGCCAGCTCGGGCAATATATGTAAGACCTGTCTCTACAAAAAATGAAAAA- | 4778 |

| | | | |
|----|------|--|------|
| QY | 1193 | AAAAAGAAAGAGAGAGGCGCGGGCGTGGCTCAGCCCTGTATCCAGCACTTTGGGAG | 1252 |
| | 4777 | -----TTGGCGAGCGCTGTGGCTCACCCCTGTATTTCAGCACTTTGGGAG | 4823 |
| Db | | | |
| QY | 1253 | GCCGAGACGGCGGAGTCACTGTGTGAGAGTTTGAGACAGCCCTGGCCAAATGGCAA | 1312 |
| | 4824 | GCCGAAAGGGGCGAGATTACCTGAGGTGAGAGTTGAGACAGCGCTGGCCAAATTGCAA | 4883 |
| Db | | | |
| QY | 1313 | AACCCCGTCTGTACTCAAAATGCAAAATTAGCCAGGCGTGGTAGCAGGCACTGTATC | 1372 |
| | 4884 | AAOCCCGGTCTACTAAAAATATACAAATATTAGCTGGGCGTGGTGCGCACCTGTATTC | 4943 |
| Db | | | |
| QY | 1373 | CCAGCTACTTGGAGAGGCTGAGGCGAGAGATTCGCTTGAACCCAGAGAGTGAAGTTGAG | 1433 |
| | 4944 | CCAGCTACTTGGAGAGGCTGAGGCGAGAGAAATCACTTGAATCTGGAGGTTGAGGTTGAG | 5003 |
| Db | | | |
| QY | 1433 | TAAAGCTGAGATGTGCGCGTTGCACTCCAGGCTGGGCGACAAGAGCAAGACTGTCTCAG | 1492 |
| | 5004 | TGAGCGGAGATGCAACCTTACACTACAGCCTGGGCGAAGAGGTGAATCTGTCTCAA | 5065 |
| Db | | | |
| QY | 1493 | AAAAAAAAAAAAAAAAAGAGAGAGAGAGAAAGGAAACA | 1551 |
| | 5064 | AAATAAATATAAATATAAATATAAATATAAATATATTA | 5102 |
| Db | | | |

```

RESULT 14
US-10-242-515-2857
Sequence 2857, Application US/10242515
Publication No. US20040009488A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC005051
CURRENT APPLICATION NUMBER: US/10/242,515
CURRENT FILING DATE: 2002-09-13
PRIORITY APPLICATION NUMBER: 09/764,877
PRIORITY FILING DATE: 2001-01-17
PRIORITY APPLICATION NUMBER: 60/119,065
PRIORITY FILING DATE: 2000-01-31
PRIORITY APPLICATION NUMBER: 60/180,628
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: 60/214,886
PRIORITY FILING DATE: 2000-06-28
PRIORITY APPLICATION NUMBER: 60/217,487
PRIORITY FILING DATE: 2000-07-11
PRIORITY APPLICATION NUMBER: 60/225,758
PRIORITY FILING DATE: 2000-08-14
PRIORITY APPLICATION NUMBER: 60/220,963
PRIORITY FILING DATE: 2000-07-26
PRIORITY APPLICATION NUMBER: 60/217,496
PRIORITY FILING DATE: 2000-07-11
PRIORITY APPLICATION NUMBER: 60/225,447
PRIORITY FILING DATE: 2000-08-14
PRIORITY APPLICATION NUMBER: 60/218,290
PRIORITY FILING DATE: 2000-07-14
Remaining SEQ ID NOS: 4031
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2857
LENGTH: 11821
TYPE: DNA
ORGANISM: Homo sapiens
US-10-242-515-2857

```

| | | | | |
|---------------------------|--------|--------------------|------------|---------------|
| Query Match | 12.9% | Score 280.2; | DB 15; | length 11821; |
| Best Local Similarity | 74.4%; | Pred. No. 4.5e-48; | | |
| Matches 386; Conservative | 0; | Mismatches 118; | Indels 15; | Gaps 2; |

QY 1013 AGGTATTTACTATATATAAGAGCTGTATACAAACAGATGATAAAGCTGCGAACCGG 1072
 Db 4599 AGGAACCTGCAAACTATGTTCAGACGACAGAGAAATGATACATGAAAGCCAGGCTGATG 4658

| | | | |
|----|------|---|------|
| QY | 1073 | GCAACGCGTCAATAGTTTACGCTGCTGGGAGGTTGAGAGGAGGAGATCGCTTGAACAAG | 1132 |
| Db | 4659 | GCTATGCGCGGTAATCCGACACCTTTAGGA-ATTTAGGCAAGAGATCACTTGAGGCCAG | 4717 |
| QY | 1133 | GTTGTTCAAGCGCGACCTGGGCAATATCAAGATCTGTCTCTGCAAAAAAAAAAAAAA | 1192 |
| Db | 4718 | GAGTTTCAAGACCAAGTCTGGGCAMATATAGTATAGACTTGTCTCTCAAAAAAAAAAAAAA-4776 | |
| QY | 1193 | AAAAAGAAAGAGAGGCGCGGGCGCTGCTGCTCAAGCTGTATATCCACGACTTTGGAG | 1252 |
| Db | 4777 | -----TTGGCAGGCGGCGGCTCAAGCCTGTATATCCACGACTTTGGAG4823 | |
| QY | 1253 | GCCGAGCGCGCGGATCACTGTGGTCAAGATTGAGACCAAGCGCTGGCCAACTGGCAA | 1312 |
| Db | 4824 | GCCGAGCGGSCAAGTTTACTGAGGTCAAGAGTTTGAGACCAAGCTGGCCAACTTGGAA4883 | |
| QY | 1313 | AACTCCGCTGTACTCAAAATGCAAAATTTGCCAGGCGCTGTGACGACACTGTATATC | 1372 |
| Db | 4884 | AACTCCGCGCTGTACTTAAAAATACAAAAATTTAGCTGGGGGTGGTGGCACGCACTGTATATC4943 | |
| QY | 1373 | CCAGCTACTTGGAGGCTGAGGACGAGAAATTCGCTTGAACCCAGAGAGGTGAGGTTGGAC | 1432 |
| Db | 4944 | CCAGCTACTTGGAAAGCTGAGGCGAGGAAATCACTTGAATCTGGAGGTTGAGGTTGCAAG5003 | |
| QY | 1433 | TAACTGAGATCGTGCCTGTGCACTCCAGCCTGGGCGCAAGAGCAAGACTCTGTCTCAG | 1492 |
| Db | 5004 | TGAGCGCGGATCGCACCATTTACATCACTACGCTGGGCAAGAGAGTGAACCTGTCTCAA5063 | |
| QY | 1493 | AAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAGAGAA1531 | |
| Db | 5064 | AATAAAAATAAAAATAAAAATAAAAATAAAAAT5102 | |

```

RESULT 15
US-10-003-295-3
; Sequence 3, Application US/10003295
; Publication No. US20020168741A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C0001183DIV
; CURRENT APPLICATION NUMBER: US/10/003,295
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-003-295-3

```

| | | | | |
|-----------------------|-----------------|--------------------|------------|---------------|
| Query Match | 12.9%; | Score 280; | DB 13; | Length 15297; |
| Best Local Similarity | 74.9%; | Pred. No. 5.7e-48; | | |
| Matches 370; | Conservative 0; | Mismatches 110; | Indels 14; | Gaps 1; |

| | | | | | |
|----|-------|---|--|-------------|-------|
| QY | 101g | ATTACATATATTAAGAGAGTGTGATAAAA | CAGAAATATATAAGGTGGAA | CCGTGGGACAC | 1078 |
| Db | 11651 | ACTCCCTTAATGCCAACCTTCCAC | CAGGAGAAATTAAGATTAACCTGGCCAGTTGCTCAC | | 11710 |
| QY | 1079 | GCTCATAGTCTTAAGCTCTTGGAGGTTGA | GGAGGAGATGAGCTTGAACA | CAGGTGTC | 1138 |
| Db | 11711 | GCTGTGATCCACGACCTTTGGAGGCTGA | GCTGGGATCACTTGAAGCCAGGAGTTCC | | 11770 |
| QY | 1139 | AAGGCCAGCTGGGCAACATPACAGATCTGTCTCTCA | AAAAAAAAAAAAAAAAAAGA | | 1198 |
| Db | 11771 | AAGATCAGCTTGGCAACACAGTGAACCTCATCTGTACAAAAAATACAAAAATA | ----- | | 11825 |
| QY | 1199 | AAGAGAGAGGGCCGGGCGTGTGGCTGACCCCTGTATCC | CACACCTTTGGAGGCGCCAG | | 1258 |
| Db | 11826 | -----GACTGGGCA | CGGTGGCTACACTGTAAATCCACACCTTTGGAGGCGCCAG | | 11876 |

| | | | |
|----|-------|--|-------|
| Qy | 1259 | CCGGGCGGAATCACTTGGTCAGAGATTGAGACAGCCTGGCCAAATGGCAAAACCCC | 1318 |
| Db | 11877 | GCAGTGGATCACTTGTGGTCAGGATTTGAGACCAAGCCAAACAATGGTAAACCCC | 11936 |
| Qy | 1319 | GTCGTATCTCAAAATGCMAAATTATGACCAGGCGTGTAGCAGGACCTGTAAATCCAGCT | 1378 |
| Db | 11937 | ATCTCTATCTAAAAATTCMAAATAATTCACGAGCATGTGTGCAGTGTCTGTAAATCCAGCT | 11936 |
| Qy | 1379 | ACTTGGGAGGCTGAGGAGAGAAATGCTGTGAACCCAGGAGGTGGAGTGTGCAGTAAAGCT | 1438 |
| Db | 11997 | ACTTGGGAGGCTGAGGTGGAGAAATTTGCTTGAACCCAGAGAGCGAGGCTGCAGTGAAGCC | 12056 |
| Qy | 1439 | GAGATCGTGGCGTTTGCACTTCAGCCCTGGGCGACAAAGGCAAGAATCTGTCTTCAGAAAAA | 1498 |
| Db | 12057 | GAGATTGTGGCACTGCATCTCCAGCCCTGGGCGACAAAGATGAAAAATCTCATCTCAAAAAAAA | 12116 |
| Qy | 1499 | AAAAAAAAAAGAGA | 1512 |
| Db | 12117 | CGAAAAATCAAAAA | 12130 |

Search completed: March 13, 2004, 05:31:15
Job time : 767.198 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 13:19:34 ; Search time 6916.2 Seconds
(without alignments)
9378.080 Million cell updates/sec

Title: US-09-966-880a-15

Perfect score: 2172
Sequence: 1 cccctgtagaggttgatga.....agtagatgataactctg 2172

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estbun:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gse_hum:*
18: em_gse_hiv:*
19: em_gse_pln:*
20: em_gse_vit:*
21: em_gse_fun:*
22: em_gse_mam:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rtd:*
26: em_gse_plg:*
27: em_gse_vrl:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| C 1 | 817.4 | 37.6 | 854 | 13 | BX464578 BX464578 |
| 2 | 781.2 | 36.0 | 918 | 13 | BX391067 BX391067 |
| 3 | 705.2 | 32.5 | 820 | 13 | BG755526 BG755526 |
| 4 | 663 | 30.5 | 664 | 12 | BG754140 BG754140 |

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| C 5 | 653.2 | 30.1 | 1201 | 9 | AL581406 AL581406 |
| C 6 | 649 | 29.9 | 920 | 13 | BX392040 BX392040 |
| C 7 | 627.6 | 28.9 | 670 | 12 | BG341546 BG341546 |
| C 8 | 626 | 28.8 | 956 | 10 | BF664355 BF664355 |
| C 9 | 615.6 | 28.3 | 1201 | 13 | BX402062 BX402062 |
| C 10 | 572.6 | 26.4 | 604 | 10 | AG978582 AG978582 |
| C 11 | 567 | 26.1 | 928 | 12 | BG398364 BG398364 |
| C 12 | 562 | 25.9 | 570 | 10 | AM504807 AM504807 |
| C 13 | 498 | 22.9 | 499 | 13 | BX283642 BX283642 |
| C 14 | 493.6 | 22.7 | 511 | 12 | BG170756 BG170756 |
| C 15 | 471 | 21.7 | 517 | 12 | BG170824 BG170824 |
| C 16 | 467.8 | 21.5 | 528 | 9 | AA954956 AA954956 |
| C 17 | 466.6 | 21.5 | 939 | 12 | BG341819 BG341819 |
| C 18 | 457 | 21.0 | 843 | 12 | BG758815 BG758815 |
| C 19 | 452.8 | 20.8 | 1140 | 10 | BF664352 BF664352 |
| C 20 | 446.2 | 20.5 | 476 | 10 | AM15547 AM15547 |
| C 21 | 443 | 20.4 | 1034 | 12 | BG755005 BG755005 |
| C 22 | 442 | 20.3 | 442 | 9 | AI016902 AI016902 |
| C 23 | 441.4 | 20.3 | 464 | 10 | AM452648 AM452648 |
| C 24 | 406.2 | 18.7 | 422 | 10 | AM253139 AM253139 |
| C 25 | 365.4 | 17.0 | 374 | 10 | AM401901 AM401901 |
| C 26 | 362 | 16.7 | 693 | 10 | BF975096 BF975096 |
| C 27 | 359.2 | 16.5 | 1201 | 13 | BX402063 BX402063 |
| C 28 | 358.2 | 16.5 | 363 | 9 | AI249134 AI249134 |
| C 29 | 353.4 | 16.3 | 413 | 10 | AM768364 AM768364 |
| C 30 | 353 | 16.3 | 353 | 9 | AA831307 AA831307 |
| C 31 | 330.4 | 15.2 | 332 | 9 | AA504653 AA504653 |
| C 32 | 321 | 14.8 | 321 | 9 | AI468242 AI468242 |
| C 33 | 317.8 | 14.6 | 364 | 10 | AM182255 AM182255 |
| C 34 | 304.6 | 14.0 | 1201 | 9 | AL559877 AL559877 |
| C 35 | 280.6 | 12.9 | 581 | 10 | AM970987 AM970987 |
| C 36 | 280.2 | 12.9 | 730 | 13 | BG616124 BG616124 |
| C 37 | 279.4 | 12.9 | 4636 | 11 | BC028737 BC028737 |
| C 38 | 277 | 12.8 | 454 | 10 | AM14750 AM14750 |
| C 39 | 276.2 | 12.7 | 822 | 14 | CB988992 CB988992 |
| C 40 | 276.2 | 12.7 | 839 | 14 | CD109867 CD109867 |
| C 41 | 275.8 | 12.7 | 705 | 28 | AQ115225 AQ115225 |
| C 42 | 274.4 | 12.6 | 769 | 13 | B0854207 B0854207 |
| C 43 | 273.4 | 12.6 | 613 | 9 | AV733538 AV733538 |
| C 44 | 273.4 | 12.6 | 696 | 13 | B0181548 B0181548 |
| C 45 | 273.2 | 12.6 | 675 | 29 | AG176894 AG176894 |

ALIGNMENTS

RESULT 1
BX464578/c 854 bp mRNA linear EST 22-MAY-2003
LOCUS BX464578 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens CDNA
DEFINITION clone CS0DG003YB14 3-PRIME, mRNA sequence.
ACCESSION BX464578
VERSION BX464578.1 GI:1029653
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Li, W.B., Gruber, C., Jeesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Genoscope

Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6672.r for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS1D0012F10NP1&cluster=6672.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CSIDG001ZF10NP1.
Location/Qualifiers
1. 854
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D003YB14"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 37.6%; Score 817.4; DB 13; Length 854;
Best Local Similarity 96.1%; Pred. No. 5.5e-110;
Matches 816; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

```

QY 307 GAACGACAGACCTGGAGACATCTAAAGTGTCAAGCTTTTCTATGACCTTTAGTAGG 366
DB 854 GAACGACAGACCTGGAGACATCTAAAGTGTCAAGCTTTTCTATGACCTTTAGTAGG 795
QY 367 ATGAGAGCAGAGGATGATCTTAAAGATGTGAGAGATCAAACTGTTTATATCA 426
DB 794 ATGAGAGCAGAGGATGATCTTAAAGATGTGAGAGATCAAACTGTTTATATCA 735
QY 427 CACCCCTTATTTTATGATTCATTTAGTAAAGTGTGATGATGATTTTCTAT 486
DB 734 CACCCCTTATTTTATGATTCATTTAGTAAAGTGTGATGATGATTTTCTAT 675
QY 487 CTTTCCCTTGAAGTTTCTTCAAGTAAACAGAACTCTTCATCAGGCAATGATCTA 546
DB 674 CTTTCCCTTGAAGTTTCTTCAAGTAAACAGAACTCTTCATCAGGCAATGATCTA 615
QY 547 GGAACCTCTAATGAGATGTCGGGTGATTTGAGACCCCAACCAATCTCTCAAGCATTA 606
DB 614 GGAACCTCTAATGAGATGTCGGGTGATTTGAGACCCCAACCAATCTCTCAAGCATTA 555
QY 607 ATATCCATATGCGCTGTATGTTTAAACAGAGAGATTTTAAATGTTTGTACAA 666
DB 554 ATATCCATATGCGCTGTATGTTTAAACAGAGAGATTTTAAATGTTTGTACAA 495
QY 667 AGAAGATTTTATGGGTGGAGATGAGATGATGACATGATGATCACTTCAAGTACT 726
DB 494 AGAAGATTTTATGGGTGGAGATGAGATGATGACATGATGATCACTTCAAGTACT 435
QY 727 TTAATAAGGATCTTAAATGAGGACAGAGACTGTGAAACAGAACCTTAATATGAGTT 786
DB 434 TTAATAAGGATCTTAAATGAGGACAGAGACTGTGAAACAGAACCTTAATATGAGTT 375
QY 787 GATGCTGAAGTACAAATCTTGTGAAACGCAACTCTTTTAAAGAACTCCTTAATTTA 846
DB 374 GATGCTGAAGTACAAATCTTGTGAAACGCAACTCTTTTAAAGAACTCCTTAATTTA 315
QY 847 GAAACCCCAACAACTTCAATCATATTAAGCAACAATGGAAGAGTGTGTTGA 906
DB 314 GAAACCCCAACAACTTCAATCATATTAAGCAACAATGGAAGAGTGTGTTGA 255
QY 907 TGTGGGAGAGAGAAATCTATTGCTCTGTGGGTCTCTGATCTCAGAAATGCAATC 966
DB 254 TGTGGGAGAGAGAAATCTATTGCTCTGTGGGTCTCTGATCTCAGAAATGCAATC 195
QY 967 AGGTCAAGTTTCTACATTTTGTATGTGTGATGCTTCTCCAAAGGTATATTA 1026
DB 194 AGGTCAAGTTTCTACATTTTGTATGTGTGATGCTTCTCCAAAGGTATATTA 135
QY 1027 TATAAGAGATTTGACAAACAGAAATGATTAAGTGTGCAACCGTGGCAACAGTCTA 1086
DB 134 TATAAGAGATTTGACAAACAGAAATGATTAAGTGTGCAACCGTGGCAACAGTCTA 75

```

QY 1087 TTCTAGCTCTTGGAGAGTGTGAGAGAGATGCTTGAACACAGGTGTTCAAGCCAG 1146
DB 74 TTCTAGCTCTTGGAGAGTGTGAGAGAGATGCTTGAACACAGGTGTTCAAGCCAG 15

QY 1147 CCTGGGCA 1155
DB 14 CCTGGGCA 6

RESULT 2

EX391067 918 bp mRNA linear EST 13-MAY-2003
LOCUS EX391067 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL012YD18 5-PRIME, mRNA sequence.
ACCESSION EX391067
VERSION EX391067.1 GI:30615325
KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 918)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6672.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BA9061ZH04_CS05792_1&cluster=6672.r.
Contact : Feng Liang Email : fliang@life.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BA9061ZH04_CS05792_1.

FEATURES

Source

Location/Qualifiers
1. 918
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL012YD18"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 36.0%; Score 781.2; DB 13; Length 918;
Best Local Similarity 95.0%; Pred. No. 9.9e-105;
Matches 872; Conservative 0; Mismatches 38; Indels 8; Gaps 6;

```

QY 134 TCTGTTTATTTCTTCACTCTCACTTTTATAGTTTA-CAGAAAAATATTATATA-191
DB 2 TTTTATTTTCTTAACCTTCACTTTTCTTAAAGTTTACGAAAAAATATTATATAC 61
QY 192 CGACTCTTAAAGATCTATGCTTGAAGAAATGAGAGAAACACAGGTCTGCCAGGGA 251
DB 62 CGACTCTTAAAGATCTATGCTTGAAGAAATGAGAGAAACACAGGTCTGCCAGGGA 121
QY 252 CGTCTGCAATTTGTTGAGATTTGAATGCAACTTCCCTACTGGGAATACAGACT 311
DB 122 CGTCTGCAATTTGTTGAGATTTGAATGCAACTTCCCTACTGGGAATACAGACT 161
QY 312 GCAGACCTGGAGACATCTTAAGTGTCAAGCTTTTCTATGACTTTTATAGTATGAG 371
DB 182 GCAGACCTGGAGACATCTTAAGTGTCAAGCTTTTCTATGACTTTTATAGTATGAG 241

```


QY 372 AGCAGAGGTAGATCTCTAAAGATGAGAGATCAATGTTTATATCAATCC 431
 Db 242 AGCAGAGGTAGATCTCTAAAGATGAGAGATCAATGTTTATATCAATCC 301
 QY 432 TTTATATTTGATTCATTTGAGTTAAAGTGTGTGATGATTTTCTATTTCTTT 491
 Db 302 TTTATATTTGATTCATTTGAGTTAAAGTGTGTGATGATTTTCTATTTCTTT 361
 QY 492 CCGTTGACGTTTCTTTCAAGTACCAACTCTTCATCAGGCGATCTATPAGGAC 551
 Db 362 CCGTTGACGTTTCTTTCAAGTACCAACTCTTCATCAGGCGATCTATPAGGAC 421
 QY 552 TCCTATGAGAGATCTGGGTGATGTGACCCCAACATCTCTCCAAAGCATTAATATC 611
 Db 422 TCCTATGAGAGATCTGGGTGATGTGACCCCAACATCTCTCCAAAGCATTAATATC 481
 QY 612 CAATCATGCGCTGTATGTTTAAATCAGCAGAGAGATGTTTATGTTTGACAAAGAG 671
 Db 482 CAATCATGCGCTGTATGTTTAAATCAGCAGAGAGATGTTTATGTTTGACAAAGAG 541
 QY 672 ATTGTTATGGGTGGG--GATGAGGTATAGACCATGATGTCACCTTCAAGTACTTTA 729
 Db 542 ATTGTTATGGGTGGGAGTGGGAGGTATAGACCATGATGTCACCTTCAAGTACTTTA 601
 QY 730 ATTAAGGATCTTAATAGGCGAGAGAGATGTAACAACACCCCTAATATGAGGTAT 789
 Db 602 ATTAAGGATCTTAATAGGCGAGAGAGATGTAACAACACCCCTAATATGAGGTAT 661
 QY 790 GTCTGAAGTACCAATCTTCTGGAACGCAAACTCTTTAAGGAATCCCTAATTAGAA 849
 Db 662 GTCTGAAGTACCAATCTTCTGGAACGCAAACTCTTTAAGGAATCCCTAATTAGAA 721
 QY 850 ACAACCAACAACTTCATCATATCATTAATAGCAAAACATTTG--AAGAAAGTCTTGATG 908
 Db 722 ACAACCAACAACTTCATCATATCATTAATAGCAAAACATTTG--AAGAAAGTCTTGATG 781
 QY 909 TTGGGAGAGAGAAATCTATTTGCTGCTGCTGCTCTTCAATCTCAAGAAATCCATCAG 968
 Db 782 TTGGGAGAGAGAAATCTATTTGCTGCTGCTGCTCTTCAATCTTCAAGAAATCCATCAG 839
 QY 969 GTCAAGGTTGTCACATTTGTTGATGTGATGATGCTTCCCAAGGTAATTAATATA 1028
 Db 840 GTCAAGGTTGTCACATTTGTTGATGTGATGATGCTTCCCAAGGTAATTAATATA 898
 QY 1029 TAAGAGATTGTGACAAA 1046
 Db 899 TAAGAGATTGTGACACA 916

RESULT 3
 BG755526 820 bp mRNA linear EST 15-MAY-2001
 LOCUS 602716206F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4856568 5',
 DEFINITION mRNA sequence.
 BG755526
 BG755526.1 GI:14066179
 EST.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strauberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Louis M. Straub, M.D., Ph.D.
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:

FEATURES
 source
 1..820
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4856568"
 /issue_type="primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_48"
 /note="Organ: B-cells; Vector: pOTB7; Site: 1; XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGACGGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 32.5%; Score 705.2; DB 12; Length 820;
 Best Local Similarity 94.5%; Pred. No. 1.3e-93;
 Matches 775; Conservative 0; Mismatches 38; Indels 7; Gaps 4;

QY 12 GGTGATGACTTACGACGCGATTCGTAATTTGGACCTTGAATGCAATCTCCAGAAAT 71
 Db 2 GGTGATGACTTACGACGCGATTCGTAATTTGGACCTTGAATGCAATCTCCAGAAAT 60
 QY 72 GTCAACACGATGAAATCTCTGCTGAAGACAGTGAATAAACAGTCTTCAAGTCT 131
 Db 61 GTCAACACGATGAAATCTCTGCTGAAGACAGTGAATAAACAGTCTTCAAGTCT 120
 QY 132 TCTGCTTTTATCTTCAACTCTCACTTCTTGAATTTACAGAAAATATTATATA 191
 Db 121 TCTGCTTTTATCTTCAACTCTCACTTCTTGAATTTACAGAAAATATTATATA 180
 QY 192 CGACTCTTTAAAGATCTATGCTTGAATAAGAGAGACACAGTCTGCCAGGGA 251
 Db 181 CGACTCTTTAAAGATCTATGCTTGAATAAGAGAGACACAGTCTGCCAGGGA 240
 QY 252 CGTGTGCAATTTGTCAGTTTGAATGCAATTTGCCCTACTGGGAATTAACAAGACT 311
 Db 241 CGTGTGCAATTTGTCAGTTTGAATGCAATTTGCCCTACTGGGAATTAACAAGACT 300
 QY 312 GCAGGACCTGGGACATCTTAAAGTGCAGGCTTTCTATGACTTTAGTAGATGAG 371
 Db 301 GCAGGACCTGGGACATCTTAAAGTGCAGGCTTTCTATGACTTTAGTAGATGAG 360
 QY 372 AGCAGAGGTAGATCTCTAAAGATGAGAGATCAATGTTTATATCAATCC 431
 Db 361 AGCAGAGGTAGATCTCTAAAGATGAGAGATCAATGTTTATATCAATCC 420
 QY 432 TTTATATTTGATTCATTTGAGTTAAAGTGTGTGATGATTTTCTATTTCTTT 491
 Db 421 TTTATATTTGATTCATTTGAGTTAAAGTGTGTGATGATTTTCTATTTCTTT 480
 QY 492 CCGTTGACGTTTCTTTCAAGTACCAACTCTTCATCAGGCGATCTATPAGGAC 551
 Db 481 CCGTTGACGTTTCTTTCAAGTACCAACTCTTCATCAGGCGATCTATPAGGAC 540
 QY 552 TCCTATGAGAGATCTGGGTGATGTGACCCCAACATCTCTCCAAAGCATTAATATC 611
 Db 541 TCCTATGAGAGATCTGGGTGATGTGACCCCAACATCTCTCCAAAGCATTAATATC 600
 QY 612 CAATCATGCGCTGTATGTTTAAATCAGCAGAGAGATGTTTATGTTTGACAAAGAG 671
 Db 601 CAATCATGCGCTGTATGTTTAAATCAGCAGAGAGATGTTTATGTTTGACAAAGAG 660
 QY 672 ATTGTTATGGGTGGGAGTAT--AGACCATGATGTCACCTTCAAG--CTACT 726

Db 661 ATTGTCATGGTGGGATGCGCCGCTCAGACCATGATGCTCACTTCAAGCTACTT 720
 Qy 727 TTAATTAAGGATCTTAAATGCGCAGAGAGACTGTGAAACAGACCCCTAATAGGTT 786
 Db 721 TTAATTAAGGATCTCAGATGGGCGACGAGACTGTGAAACAGAAATCTTAATATGGGT 780
 Qy 787 GATGCTGAAG-TAGCAATCTTCTGGAACGCAACTCT 825
 Db 781 GATGCTGAAGTTAGCAATCTCTGGAACGCAACT 820

RESULT 4
 BG754140 664 bp mRNA linear EST 15-MAY-2001
 LOCUS 602709681F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4846005 5',
 DEFINITION mRNA sequence.
 ACCESSION BG754140
 VERSION BG754140.1 GI:14064793
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/
 TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: csaps-r@mail.nih.gov
 Tissue Procurement: Louis M. Straud, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNCM1686 row: b column: 22
 High quality sequence stop: 662.
 Location/Qualifiers

FEATURES
 source 1..664
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4846005"
 /issue_type="primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH_MGC_48"
 /note="Organ: B-cells; Vector: pOT7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 30.5%; Score 663; DB 12; Length 664;
 Best Local Similarity 100.0%; Freq. No. 2.1e-87;
 Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 12 GGTGATGAGCTTATGAGAGAGCAATTCGACTTGGGACTTGTAGCACTTCAGAAAT 71
 Db 2 GGTGATGAGCTTATGAGAGAGCAATTCGACTTGGGACTTGTAGCACTTCAGAAAT 61
 Qy 72 GTCAACACGATGAATATCTCTGCTGAGAGCAGTGATGATGATGATGATGATGATGAT 131
 Db 62 GTCAACACGATGAATATCTCTGCTGAGAGCAGTGATGATGATGATGATGATGATGAT 121
 Qy 132 TCTGTGTTTATTTCTTCAACTCTCACTTCTTGAAGTTTACAGAAAAATATTTATTA 191
 Db 122 TCTGTGTTTATTTCTTCAACTCTCACTTCTTGAAGTTTACAGAAAAATATTTATTA 181

Qy 192 CGACTCTTTAAAGATCTATGCTTGAATAAGAGAGAAACAGAGTCTGGCAGGGA 251
 Db 182 CGACTCTTTAAAGATCTATGCTTGAATAAGAGAGAAACAGAGTCTGGCAGGGA 241
 Qy 252 CGTGTGCAATTTGGTCAAGTTTGAATGCAATTGTCCCTTACTGGGAATTAAGAACT 311
 Db 242 CGTGTGCAATTTGGTCAAGTTTGAATGCAATTGTCCCTTACTGGGAATTAAGAACT 301
 Qy 312 GCAGGACTTGGAGCATCTTAAGTGTCAAGCTTTTCTATGACTTTAGTAGATGAG 371
 Db 302 GCAGGACTTGGAGCATCTTAAGTGTCAAGCTTTTCTATGACTTTAGTAGATGAG 361
 Qy 372 AGCAGAGATGATGCTTAAAGACATGATGAGAGATCAAAATGTTTATCAACATCC 431
 Db 362 AGCAGAGATGATGCTTAAAGACATGATGAGAGATCAAAATGTTTATCAACATCC 421
 Qy 432 TTAATTAATTTGATTCATTTGATGATTAAGTGTGTTAGTATGATTTTCTATTTT 491
 Db 422 TTAATTAATTTGATTCATTTGATGATTAAGTGTGTTAGTATGATTTTCTATTTT 481
 Qy 492 CCTTGACGTTTACTTTCAGATACAGCAAACTCTTCCATCAGGCCATGATAGAGCC 551
 Db 482 CCTTGACGTTTACTTTCAGATACAGCAAACTCTTCCATCAGGCCATGATAGAGCC 541
 Qy 552 TCCCTAATGAGATGATGCTGGTGTGATGAGCCCAACCATCTCTCCAAAGCATTAATTC 611
 Db 542 TCCCTAATGAGATGATGCTGGTGTGATGAGCCCAACCATCTCTCCAAAGCATTAATTC 601
 Qy 612 CAATCATGCGCTGATGTTTAAATCAAGCAGAGCATGTTTATGTTGTGACAAAAGAG 671
 Db 602 CAATCATGCGCTGATGTTTAAATCAAGCAGAGCATGTTTATGTTGTGACAAAAGAG 661
 Qy 672 ATT 674
 Db 662 ATT 664

RESULT 5
 AL581406/c 1201 bp mRNA linear EST 01-JUN-2003
 LOCUS AL581406 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
 DEFINITION clone CS0DG003YB14 3-PRIME, mRNA sequence.
 ACCESSION AL581406
 VERSION AL581406.2 GI:31319646
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
 1 (bases 1 to 1201)
 TITLES Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On Feb 16, 2001 this sequence version replaced gi:12948375.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 6672.r For more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DG003DA07NP1&cluster=6672.r. Contact :
 Feng Liang Email : fliang@life.techn.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DG003DA07NP1.

FEATURES

source 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DG003YB14"

/issue_type="B CELLS (RAMOS CELL LINE)"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

Query Match 30.1%; Score 653.2; DB 9; Length 1201;
 Best Local Similarity 89.8%; Pred. No. 3.8e-86;

Matches 680; Conservative 31; Mismatches 42; Indels 4; Gaps 2;

QY 404 AGAGTCAAGTGTATATATCAACACCTTTATTTATGATTCATTGATTAACAGTG 463
 DB 756 AADATATCAAAATTTTATATATCAAAATCTTTATTTATTTATTCATTATTA-TTAACATTGT 698
 QY 464 TGTATGATAGATTTTCTATCTTTCCCTGAGTTTACTTCAAGTAAACAACACT 523
 DB 697 TTTATATATATATATTTTCTATTTTCTGCTGAGGTTTACTTCAAAATTAAGAAATT 638
 QY 524 CTTCATCAAGGCGATGATCTATAGACCTCTATAGAGATCTGGGTGATTGTGACC 583
 DB 637 CTTSSTTSAKGCATATCTATAGGCGCTTATAGAGATNCTGGGTGATTGTGACVC 578
 QY 584 CAACCACTCTTCCAAAGCATTTATATCCATATCAGCGCTGATGTTTATACAGAGAA 643
 DB 577 CAACCAACNCTCCAAAGCATTTATATCCATATCAGCGCTGATGTTTATACAGAGAA 518
 QY 644 GCATGTTTTATGTTGTGCAAAAGAGTGTATAGGTTGGGATGAGATATAGACA 703
 DB 517 GCATGTTTTATGTTGTGCAAAAGAGTGTATAGGTTGGGATGAGATATAGACA 458
 QY 704 TGCATGTCACCTTCAAGCTACTTTATTAAGATCTTAAATGGGAGAGAGACTGTA 763
 DB 457 TGCATGTCACCTTCAAGCTACTTTATTAAGATCTTAAATGGGAGAGAGACTGTA 398
 QY 764 ACAAGACACCTTATATATGAGTGTATGATGATGAGCAATCTTCTGGAACGCAACT 823
 DB 397 ACAAGACACCTTATATATGAGTGTATGATGATGAGCAATCTTCTGGAACGCAACT 338
 QY 824 CTTTAAAGAGTCCCTTAATTTAGAAACCCCAACACTTCAATATCATATTTGCAAA 883
 DB 337 CTTTAAAGAGTCCCTTAATTTAGAAACCCCAACACTTCAATATCATATTTGCAAA 278
 QY 884 CAATTGAGAGAGTGTGTAATGTTGGGAGAGAGAAATCTATTGCTCTCGGGTTC 943
 DB 277 CAATTGAGAGAGTGTGTAATGTTGGGAGAGAGAAATCTATTGCTCTCGGGTTC 218
 QY 944 TCTTCATCTCAGAAATGCAATCAGGTCAAGGTTTGCTCATTGTTATGTGTGATGC 1003
 DB 217 TCTTCATCTCAGAAATGCAATCAGGTCAAGGTTTGCTCATTGTTATGTGTGATGC 158
 QY 1004 TTTCTCCCAAGGTATATTAATATATATAGAGTGTGCAAAACAGAAATGATAAGCTG 1063
 DB 157 TTTCTCCCAAGGTATATTAATATATATAGAGTGTGCAAAACAGAAATGATAAGCTG 98
 QY 1064 CGAACCGTGGACACCTCATAGTTTACTGTTGGGAGGTTGAGAGAGAGATGGCT 1123
 DB 97 CGAACCGTGGACACCTCATAGTTTACTGTTGGGAGGTTGAGAGAGAGATGGCT 38
 QY 1124 TGAACACAGGTGTTCAAGGCGAGCTGGGCAACATTA 1160
 DB 37 ---GAAACAAGTGTCAAGGCGAGCTGGGCAACATTA 4

RESULT 6
 BX392040/c 920 bp mRNA linear EST 13-MAY-2003
 LOCUS BX392040 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 DEFINITION Homo sapiens CDNA clone CS0DL012YD18 3-PRIME, mRNA sequence.
 ACCESSION BX392040

VERSION BX392040.1 GI:30607824
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 920)
 AUTHORS Li, W. B., Gruber, C., Jesse, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by life technologies, a division of
 invitrogen. This sequence belongs to sequence cluster 6672.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0BA1048ZH08_CS04576_1cluster=6672.r.
 Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0BA1048ZH08_CS04576_1.
 Location/Qualifiers
 1..920
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DL012YD18"
 /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /clone_lib="RAMOS CELL LINE"
 /issue_type="Homo sapiens B CELLS (RAMOS CELL LINE) COT
 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoRV
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 29.9%; Score 649; DB 13; Length 920;
 Best Local Similarity 98.5%; Pred. No. 1.8e-85;
 Matches 655; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CCCTGTATAGAGTGTATGATCTTACGAGAGCAATTCCTGTTGGACTTATAGCA 60
 DB 665 CCCCTGTATAGAGTGTATGATCTTACGAGAGCAATTCCTGTTGGACTTATAGCA 606
 QY 61 CTTCCAGAAATGTCACACAGATGAATATCTGCTGGAAGACAGTGAATAAAACAGT 120
 DB 605 CTTCCAGAAATGTCACACAGATGAATATCTGCTGGAAGACAGTGAATAAAACAGT 546
 QY 121 CTTCAAGTCTTCTGTTTATTTATTTCAACTCTCACTTTCTAGAGTTTACAGAAAA 180
 DB 545 CTTCAAGTCTTCTGTTTATTTATTTCAACTCTCACTTTCTAGAGTTTACAGAAAA 486
 QY 181 ATATTATATACGACTTTTAAAGATCTATGCTTGAATAATAGAGAGAAACAAGT 240
 DB 485 ATATTATATACGACTTTTAAAGATCTATGCTTGAATAATAGAGAGAAACAAGT 426
 QY 241 CTGGCCAGGAGACGTCGCAATTTGGGCAATTTTGAATGGAATGTCCTCCCACTGGGA 300
 DB 425 CTGGCCAGGAGACGTCGCAATTTGGGCAATTTTGAATGGAATGTCCTCCCACTGGGA 366
 QY 301 ATAAACAGACTGACGACCTGGAGCATCTTAAGTGTCAACGTTTTTCTATGACTTTTA 360
 DB 365 ATAAACAGACTGACGACCTGGAGCATCTTAAGTGTCAACGTTTTTCTATGACTTTTA 306
 QY 361 GGTAGATGAGAGAGAGTAGATCTTAAAGACATGTTGAGAGATCAAAATGTTTTTA 420
 DB 305 GGTAGATGAGAGAGAGTAGATCTTAAAGACATGTTGAGAGATCAAAATGTTTTTA 246
 QY 421 TATCAACATCTTTATTTATTTGATTCATTTGAGTTAAAGGTGTTTATGATAGATTTT 480
 DB 245 TATCAACATCTTTATTTATTTGATTCATTTGAGTTAAAGGTGTTTATGATAGATTTT 186

QY 481 TCTATCTTCCCTGAGCTTAACTTCAAGTAACACAACTCTTCATCAGGCGATGA 540
 DB 185 TCTATCTTCCCTGAGCTTAACTTCAAGTAACACAACTCTTCATCAGGCGATGA 126
 QY 541 TCTATGAGCCTCTTAATGAGAGTATCTGGGTGATTTGAGCCCAACATCTCTCCAA 600
 DB 125 TCTATGAGCCTCTTAATGAGAGTATCTGGGTGATTTGAGCCCAACATCTCTCCAA 66
 QY 601 GCATTATATCCATATGCGCTGATGTTTAAATCAGAGAGAGATGTTTATGTTG 660
 DB 65 GCATTATATCCATATGCGCTGATGTTTAAATCAGAGAGAGATGTTTATGTTG 6
 QY 661 TACA 665
 DB 5 AAAAA 1

RESULT 7
 LOCUS BG341546 670 bp mRNA linear EST 27-FEB-2001
 DEFINITION 602463652P1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4576548 5',
 mRNA sequence.
 ACCESSION BG341546
 VERSION BG341546.1 GI:13147984
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 670)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rcmail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LNCMI289 row: g column: 13
 High quality sequence stop: 650.
 Location/Qualifiers
 1. 670
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4576548"
 /rissue_type="primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_48"
 /note="Organ: B-cells; Vector: pOT8; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 28.8%; Score 627.6; DB 12; Length 670;
 Best Local Similarity 98.8%; Pred. No. 3e-82;
 Matches 664; Conservative 0; Mismatches 4; Indels 4; Gaps 3;

QY 12 GGTGATGACTTACGAGACGATTCGTAATGGAGCTTGTAGCACTTCAGGAAT 71
 DB 2 GGTGATGACTTACGAGACGATTCGTAATGGAGCTTGTAGCACTTCAGGAAT 60

QY 72 GTCACACAGATGAATATCTCTGCTGAAGACAGATGATAAAAACAGTCTTCAAGTCT 131
 DB 61 GTCACACAGATGAATATCTCTGCTGAAGACAGATGATAAAAACAGTCTTCAAGTCT 120
 QY 132 TCTCTGTTTATCTTCAACTCTCTTCTTAGAGTTTACAGAAAAATTTATATA 191
 DB 121 TCTCTGTTTATCTTCAACTCTCTTCTTAGAGTTTACAGAAAAATTTATATA 180
 QY 192 CGACTCTTTAAAGATCTATGCTTGAAGAAATGAGAGGAAACAGAGTCTGGCCAGGA 251
 DB 181 CGACTCTTTAAAGATCTATGCTTGAAGAAATGAGAGGAAACAGAGTCTGGCCAGGA 240
 QY 252 CGTGTGCAATTTGATGAGTTTGAATGCAACATGTCCTTACTGGGAATTAACAAGT 311
 DB 241 CGTGTGCAATTTGATGAGTTTGAATGCAACATGTCCTTACTGGGAATTAACAAGT 300
 QY 312 GCAGGACCTGGAGCATCTTAAAGTGTCAACGTTTCTATAGCTTTAGTAGATGAG 371
 DB 301 GCAGGACCTGGAGCATCTTAAAGTGTCAACGTTTCTATAGCTTTAGTAGATGAG 360
 QY 372 AGCAGAGGTAGATCTTAAAGATGTGAGAGATCAAAATGTTTATATCAACATCC 431
 DB 361 AGCAGAGGTAGATCTTAAAGATGTGAGAGATCAAAATGTTTATATCAACATCC 420
 QY 432 TTTATATTTGATTCATTTGATTAACAGTGTGTTAGTATGATTTTCTATCTTT 491
 DB 421 TTTATATTTGATTCATTTGATTAACAGTGTGTTAGTATGATTTTCTATCTTT 480
 QY 492 CCCTTGACGTTTACTTGAAGTAACAAACCTTTCATGAGCCATGATCTATAGACC 551
 DB 481 CCCTTGACGTTTACTTGAAGTAACAAACCTTTCATGAGCCATGATCTATAGACC 540
 QY 552 TCCATGAAGATATCTGGGTGATTTGAGACCCCAACATCTCTCC-AAAGCTATATAT 610
 DB 541 TCCATGAAGATATCTGGGTGATTTGAGACCCCAACATCTCTCCAAAGCATTTATAT 600
 QY 611 CCATCATGCGCTGATGTTTATATCAGCAGAAAGCATGTTTATGTTGTACAAAAGAA 670
 DB 601 CCATCATGCGCTGATGTTTATATCAGCAGAAAGCATGTTT--AGTGTATCAAAAGAA 658
 QY 671 GATTGTTATGGG 682
 DB 659 GATTGTTATGGG 670

RESULT 8
 BFE64355 956 bp mRNA linear EST 21-DEC-2000
 LOCUS 602146013P1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4309707 5',
 mRNA sequence.
 ACCESSION BFE64355
 VERSION BFE64355.1 GI:11938351
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 956)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rcmail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LNCMI183 row: 1 column: 04
 High quality sequence stop: 725.
 Location/Qualifiers

FEATURES

Source

1. .956
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4309707"
 /cissue_type="Primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="XIH_MGC 48"
 /note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."

ORIGIN

Query Match 28.8%; Score 626; DB 10; Length 956;
 Best Local Similarity 97.2%; Pred. No. 4.1e-82;
 Matches 690; Conservative 0; Mismatches 15; Indels 5; Gaps 5;

QY 12 GGTGATGACTTACGAGACGATTTGTACTTTGGACTTTGATAGCACTTCCAGGAT 71
 DB 2 GGTGATGACTTACGAGACGATTTGTACTTTGGACTTTGATAGCACTTCCAGGAT 60
 QY 72 GTCAACACGATGAATATCTCTGCTGAAGACAGTGATTAATAACAGTCTTCAA-GTC 130
 DB 61 GTCAACACGATGAATATCTCTGCTGAAGACAGTGATTAATAACAGTCTTCAAAGTTC 120
 QY 131 TTCTCTGTTTATCTTCAACTCTGATCTTTTAAGTTTACAGAAAAATTTATAT 190
 DB 121 TTCTCTGTTTATCTTCAACTCTGATCTTTTAAGTTTACAGAAAAATTTATAT 180
 QY 191 ACCGACTTTAAAGATCTATGCTTGAATAAGAGAGAAACAGGCTGCGCAGGG 250
 DB 181 ACCGACTTTAAAGATCTATGCTTGAATAAGAGAGAAACAGGCTGCGCAGGG 240
 QY 251 ACCGCTGCAATGTGCGAGTTTGAATGCAATGTCCTCCCTACTGGGATTAAGAC 310
 DB 241 ACCGCTGCAATGTGCGAGTTTGAATGCAATGTCCTCCCTACTGGGATTAAGAC 300
 QY 311 TGACGACCTGGAGACATCTTAAGTCAAGCTTTTCTATGACTTTTGGTAGATGA 370
 DB 301 TGACGACCTGGAGACATCTTAAGTCAAGCTTTTCTATGACTTTTGGTAGATGA 360
 QY 371 GACGACAGGTAGATCTTAAGATGAGATGAGATCAAAATTTTATATCAACATC 430
 DB 361 GACGACAGGTAGATCTTAAGATGAGATGAGATCAAAATTTTATATCAACATC 420
 QY 431 CTTATATTTTGAATCATT-TGAGTTAACAGTGTGTTAGTAGATTTTCTATCTT 489
 DB 421 CTTATATTTTGAATCATTGTGAGTTAACAGTGTGTTAGTAGATTTTCTATCTT 480
 QY 490 TTCCCTGACGTTTACTTTCAGTAAACAACTTTCATCAGGCCATGATCTATAGA 549
 DB 481 TTCCCTGACGTTTACTTTCAGTAAACAACTTTCATCAGGCCATGATCTATAGA 540
 QY 550 CCTCTCTATGAGAGATCTGGGATGATGAGACCCCAACATCTTCCCAAGCATTAATA 609
 DB 541 CCTCTCTATGAGAGATCTGGGATGATGAGACCCCAACATCTTCCCAAGCATTAATA 599
 QY 610 TCCATCATGCGGTATGTTTATCATCAGAGAGATGTTTATGTTTGTACAAAGA 669
 DB 600 TCCATCATGCGGTATGTTTATCATCAGAGAGATGTTTATG-TTGTACAAAGA 658
 QY 670 AGATTGTATGGGTGGGATGAGATATAGACCATGATGATCACTTCA 719
 DB 659 AGATTGTATGGGTGGGATGAGATATAGACCATGATGATCACTTCA 708

RESULT 9

EX402062/c 1201 bp mRNA linear EST 13-MAY-2003
 LOCUS
 DEFINITION
 Homo sapiens cDNA clone CS0D012BD18 3-PRIME, mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 1201)
 Li, W.-B., Gruber, C., Jesse, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 6672.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0D012BD09NP1&cluster=6672.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/InvitrogenCorporation 1600
 Faraday Avenue Genoscope sequence ID : CS0D012BD09NP1.
 Location/Qualifiers

FEATURES

1. .1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D012BD18"
 /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 28.3%; Score 615.6; DB 13; Length 1201;
 Best Local Similarity 77.0%; Pred. No. 1.1e-80;
 Matches 715; Conservative 92; Mismatches 107; Indels 15; Gaps 8;

QY 204 AAGATCTATGCTTGAATAATGAGAGAACACAGGCTGCCAGGAGCGTGCAAT 263
 DB 945 AAAAAATWASWAKGTATTAATRTDKAGKACASAGTGGAGKSGSKCAAT- 887
 QY 264 GGTGACGTTTGAATGACATGTCCTCCCTACTGGGATTAACAGACTGACGACCTGG 323
 DB 886 -GKTCCKTKTGKAGCACTTGTCW-AAAAAAAAAAAAAAAAANRRMNNNTK 829
 QY 324 AGCATCTTAAGTCAAGCTTTTCTATGACTTTTNGTAGATGAGACGAGAGTAG 383
 DB 828 AKNNWCCCTAAAGTCAACCTTATCTATGACAAWAAAGTAGATKAKACAAAGATAD 769
 QY 384 ATCTTAAGAGAGAGAGATCAAAAGTTTATATCAACATCTTATATTTGA 443
 DB 768 ATCTTAAGAGAGAGAGATCAAAAGTTTATATCAACATCTTATATTTGA 709
 QY 444 TTCAATGAGTACAGTGGGT-TAGTGAATGTTTCTATCTTCTTCCCTGACGT 502
 DB 708 TWCAATTKAKTAAACATGKTGTATGAGATGATTTTCTATMTTAAACCTTGACGT 649
 QY 503 T---ACTTCAAGTACAAACCTTTCATCAGGCCATGATATAGACCTTCAATG 559
 DB 648 TCAGNCATACAGTAAACAACTTTCCAAGGCCATGATATATGAGACCTCCCAAG 589
 QY 560 AGATTATGGGTATGTGACCCCAACCATCTTCCAAAGCATTAATCCATATG 619

Db 588 AGAGTACCGAGCGATGAGCCCAACCAATTCACCAAAACAATAAATCAATCAAG 529
 Qy 620 CCCTGATGCTTTTATTC-AGCAGAGCACTTTTATGTTGTACAAAAGAGTTGTA 678
 Db 528 CCCTGATGCTTTTATTCAGCAGAGCACTTTTATGTTGTACAAAAGAGTTGTA 469
 Qy 679 TGGGTGGGATGAGGATAGACCAATGCTGCTCACTTCAAGTACTTTAATAAGAT 738
 Db 468 TGGGTGGGATGAGGATAGACCAATGCTGCTCACTTCAAGTACTTTAATAAGAT 409
 Qy 739 CTTAAATGGGAGAGAGAGCTGTGAACAACACCTTAATGAGTGTGATGTC-TGAAG 797
 Db 408 CTTAAATGGGAGAGAGAGCTGTGAACAACACCTTAATGAGTGTGATGTC-TGAAG 349
 Qy 798 TAGCAATCTTCTGGAGAGCAACTTTTAAAGAGTCCCTAATTGAAACACCCAC 857
 Db 348 TAGCAATCTTCTGGAGAGCAACTTTTAAAGAGTCCCTAATTGAAACACCCAC 289
 Qy 858 AAACCTCACATATCATATATAGCAACCAATGGAAGAAAGTGTGAATGTTGGGAGA 917
 Db 288 AAACCTCACATATCATATATAGCAACCAATGGAAGAAAGTGTGAATGTTGGGAGA 229
 Qy 918 GGAATATCTATGCTCTGCTGCTGCTCTTC-ATCTCAGAAATGCCATCA--GGTTC 971
 Db 228 GGAATATCTATGCTCTGCTGCTGCTCTTC-ATCTCAGAAATGCCATCA--GGTTC 169
 Qy 972 AAGCTTGTCTACATTTTGTATGTGTGATGCTTCTCCCAAGATATTAATTAATTA 1031
 Db 168 AAGCTTGTCTACATTTTGTATGTGTGATGCTTCTCCCAAGATATTAATTAATTA 109
 Qy 1032 GAGAGTGTGACAAACAGAAATATTAAGCTGGAACCGTGGACACCGCTCATAGTCTA 1091
 Db 108 GAGAGTGTGACAAACAGAAATATTAAGCTGGAACCGTGGACACCGCTCATAGTCTA 49
 Qy 1092 GCTGCTGGAGGAGTGTGAGGAGAGATG 1120
 Db 48 GCTGCTGGAGGAGTGTGAGGAGAGATG 20

RESULT 10 604 bp mRNA linear EST 02-JUN-2000
 AM978582/c
 LOCUS AM978582 NAGE resequences, MAGP Homo sapiens cDNA, mRNA sequence.
 DEFINITION AM978582
 ACCESSION AM978582
 VERSION AM978582.1 GI:8169850
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 604)
 Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaepard,R., Gay,C.,
 Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
 Quackenbush,J.
 Assessment of gene expression patterns in a model of colon tumor
 metastasis using a 19,200 element cDNA microarray
 Unpublished (2000)
 CONTACT: John Quackenbush
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@tigr.org
 Plate: 394
 Seq primer: Forward.
 Location/Qualifiers
 1..604
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="NAGE resequences, MAGP"
 /note="Vector: pBluescriptSkm"

TITLE
 JOURNAL
 COMMENT
 FEATURES
 source
 ORIGIN

Query Match 26.4%; Score 572.6; DB 10; Length 604;
 Best Local Similarity 97.6%; Pred. No. 3.4e-74;
 Matches 581; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 439 TTTGATTCATTTAGATTAACAGTGTGTAGTATGATTTTCTATTTTCCCTTA 498
 Db 595 TTGAATTCATTTAGATTAACAGTGTGTAGTATGATTTTCTATTTTCCCTTA 536
 Qy 499 CTTTACTTTCAAGTAAACAACCTCTCCATAGGCGATGATATAGACCTCTAT 558
 Db 535 CTTTACTTTCAAGTAAACAACCTCTCCATAGGCGATGATATAGACCTCTAT 476
 Qy 559 GAGAGTATCGGTGATTTGACCCCAACATCTCTCCAAAGCATTAATATCATAT 618
 Db 475 GAGAGTATCGGTGATTTGACCCCAACATCTCTCCAAAGCATTAATATCATAT 416
 Qy 619 GCGGTATGTTTAAATCAGCAGAGCATGTTTATGTTGTACAAAAGAGTTGTA 678
 Db 415 GCGGTATGTTTAAATCAGCAGAGCATGTTTATGTTGTACAAAAGAGTTGTA 356
 Qy 679 TGGGTGGGATGAGGATAGACCATGATGTCACCTTCAAGTCTTTAATAAGAT 738
 Db 355 TGGGTGGGATGAGGATAGACCATGATGTCACCTTCAAGTCTTTAATAAGAT 296
 Qy 739 CTTAAATGGGAGAGAGCTGTGAACAACACCTTAATGAGTGTGATGCTGAAT 798
 Db 295 CTTAAATGGGAGAGAGCTGTGAACAACACCTTAATGAGTGTGATGCTGAAT 236
 Qy 799 AGCAATCTTCTGAAACGCAACTCTTTTAAAGAGTCCCTAATTAGAAACCCACA 858
 Db 235 AGCAATCTTCTGAAACGCAACTCTTTTAAAGAGTCCCTAATTAGAAACCCACA 176
 Qy 859 AACTTCACATATCAATTAATGCAACATTTGAAGAGTGTCTTAATGTTGGGAGAG 918
 Db 175 AACTTCACATATCAATTAATGCAACATTTGAAGAGTGTCTTAATGTTGGGAGAG 116
 Qy 919 GAAATCTATTTGGTCTCCGTTGCTCTTCATCTCAAAATGCAATCAGTCAAGTT 978
 Db 115 GAAATCTATTTGGTCTCCGTTGCTCTTCATCTCAAAATGCAATCAGTCAAGTT 56
 Qy 979 GCTACATTTTGTATGTTGATGATGCTTCTCCCAAGGATATTAATATTAAGA 1033
 Db 55 GCTACATTTTGTATGTTGATGATGCTTCTCCCAAGGATATTAATATTAAGA 1

RESULT 11 928 bp mRNA linear EST 12-MAR-2001
 BG398364
 LOCUS BG398364 602439832P1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:46656 5',
 DEFINITION mRNA sequence.
 ACCESSION BG398364
 VERSION BG398364.1 GI:13291812
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 928)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Yang Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNCM284 row: 0 column: 07
 High quality sequence stop: 650.

FEATURES

Location/Qualifiers
 1..928
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:456366"
 /tissue_type="primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_48"
 /note="Organ: B-cells; Vector: pOT7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using Zap-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 26.1%; Score 567; DB 12; Length 928;
 Best Local Similarity 98.8%; Pred. No. 1.7e-73;
 Matches 614; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

```

QY 1 CCCCTGATAGGTTGATGACTTACGACGACATTCGACTTGGACCTTGATGCAA 60
DB 56 CCCCTGATAGGTTGATGACTTACGACGACATTCGACTTGGACCTTGATGCAA 114
QY 61 CTCACGAGATGTCACACAGATGAATATCTCTGCTGAAGACAGTGAATAAAAAGT 120
DB 115 CTCACGAGATGTCACACAGATGAATATCTCTGCTGAAGACAGTGAATAAAAAGT 174
QY 121 CCTTCAAGTCTCTCTGTTTATTTCTTCACTCTCACTTCTTGAAGTTACAGAAAA 180
DB 175 CCTTCAAGTCTCTCTGTTTATTTCTTCACTCTCACTTCTTGAAGTTACAGAAAA 234
QY 181 ATATTTATATGACTCTTTTAAAAAGATCTATGCTTTGAAAAATGAGAAACACAGT 240
DB 235 ATATTTATATGACTCTTTTAAAAAGATCTATGCTTTGAAAAATGAGAAACACAGT 294
QY 241 CTGGCCGAGGACGCTGCAATGCTGAGTGTGATGATGCAATGTCCTTACTGGGA 300
DB 295 CTGGCCGAGGACGCTGCAATGCTGAGTGTGATGATGCAATGTCCTTACTGGGA 353
QY 301 ATACAGAACTGACAGACCTGGAGCATCTTAAAGTGTCAACGTTTTCTATGACTTTA 360
DB 354 ATACAGAACTGACAGACCTGGAGCATCTTAAAGTGTCAACGTTTTCTATGACTTTA 413
QY 361 GGTAGATGAGACAGAGGTGATCTTAAAAAGCATGCTGAGAGATCAATG-TTTT 419
DB 414 GGTAGATGAGACAGAGGTGATCTTAAAAAGCATGCTGAGAGATCAATG-TTTT 473
QY 420 AATACCAATCTTTATTTATTTGATTTGATTTGATTTGATTTGATTTGATTT 479
DB 474 AATACCAATCTTTATTTATTTGATTTGATTTGATTTGATTTGATTTGATTT 533
QY 480 TTCTATCTTTTCCCTGAGTCTTACTTTCAAGTAAACAACCTTCCATCGAGCATG 539
DB 534 TTCTATCTTTTCCCTGAGTCTTACTTTCAAGTAAACAACCTTCCATCGAGCATG 593
QY 540 ATCTATAGACCTCTTATAGAGATATCTGGGTGATTTGACCCCAACCATCTCCAA 599
DB 594 ATCTATAGACCTCTTATAGAGATATCTGGGTGATTTGACCCCAACCATCTCCAA 652
QY 600 AGCATTATATTCATCATGCGC 622
DB 653 AGCATTATATTCATCATGCGC 675

```

RESULT 12
 AM504807 570 bp mRNA linear EST 02-MAR-2000
 LOCUS
 DEFINITION UI-HF-BNO-ain-f-05-0-UI.t1 NIH_MGC_50 Homo sapiens cDNA clone

IMAGE:3080216 5', mRNA sequence.
 AM504807
 VERSION
 AM504807.1 GI:7142474
 EST.
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 CONTACT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Eco RI site shown at the beginning of the sequence.

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: M.B. Soares Lab
 cDNA Library Arrayed by: M.B. Soares Lab
 DNA Sequencing by: M.B. Soares Lab
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:
www.bio.linnl.gov/bdip/image/image.html
 Seq primer: M13 Forward.

FEATURES

Location/Qualifiers
 1..570

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3080216"
 /tissue_type="lymph"
 /cell_type="germinal center B cells"
 /cell_line="MGC85"
 /lab_host="DH10B (LTI)"
 /clone_id="NIH_MGC_50"
 /note="Vector: pTZ19-Pac; Site 1: NotI; Site 2: Eco RI; Constructed from size fractionated cytoplasmic mRNA (3.5-4.4kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

ORIGIN

Query Match 25.9%; Score 562; DB 10; Length 570;
 Best Local Similarity 99.8%; Pred. No. 1.2e-72;
 Matches 562; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 489 TTTCCCTTGAAGTTTACTTCAAGTAAACAACCTTCCATCAGGCATGATCTATAGG 548
DB 8 TTTCCCTTGAAGTTTACTTCAAGTAAACAACCTTCCATCAGGCATGATCTATAGG 67
QY 549 AACTCTTATGAGATCTGCGTGTGATTTGAGACCCCAACCATCTTCCAAAGCATTAAT 608
DB 68 AACTCTTATGAGATCTGCGTGTGATTTGAGACCCCAACCATCTTCCAAAGCATTAAT 127
QY 609 ATCCATATGCGCTGATGTTTATATCAGAGAAAGCATGTTTATGTTTGTACAAAG 666
DB 128 ATCCATATGCGCTGATGTTTATATCAGAGAAAGCATGTTTATGTTTGTACAAAG 187
QY 669 AAGATGTTATGGGAGGAGTGAAGTATAGACCATGATGATGATGATGATGATGAT 728
DB 188 AAGATGTTATGGGAGGAGTGAAGTATAGACCATGATGATGATGATGATGATGAT 247
QY 729 AATAAGGATCTTAAATGGGACAGAGACTGTGAACAAGACACCTTAATATGAGTTGA 788
DB 248 AATAAGGATCTTAAATGGGACAGAGACTGTGAACAAGACACCTTAATATGAGTTGA 307
QY 789 TGTCTGAAGTACCAATTTTGTGAAAGCGCAATCTTTTAAAGAGTCCCTAATTTTGA 848
DB 308 TGTCTGAAGTACCAATTTTGTGAAAGCGCAATCTTTTAAAGAGTCCCTAATTTGA 367
QY 849 AACCCCAACAACCTTCAATATCAATATTTACCAACAATGGAGAGAGTTGCTTAATG 908
DB 368 AACCCCAACAACCTTCAATATCAATATTTACCAACAATGGAGAGAGTTGCTTAATG 427

```


| | | | |
|----|------|---|------|
| QY | 909 | TTGGGAGAGGAAAAATCTATGGCTCTGTGGGTCTTCAATCTCAAGAAAGCCAAATAG | 968 |
| Db | 428 | TTGGGAGAGGAAAAATCTATGGCTCTGTGGGTCTTCAATCTCAAGAAAGCCAAATAG | 487 |
| QY | 969 | GTCAAGGTTTGCTACATTTGTATGTGTGATGCTTCTCCAAAGTATATTTAACTATA | 1028 |
| Db | 488 | GTCAAGGTTTGCTACATTTGTATGTGTGATGCTTCTCCAAAGTATATTTAACTATA | 547 |
| QY | 1029 | TTAAGAGTTGTGACAAACAGA | 1051 |
| Db | 548 | TTAAGAGTTGTGACAAACAGA | 570 |

| RESULT 13 | LOCUS | DEFINITION |
|-----------|--------|-------------------------------|
| EX283642 | 499 bp | mRNA |
| | | linear |
| | | EST 05-MAR-2003 |
| | | IMGP558K011707 ; |
| | | IMAGE:4856568, mRNA sequence. |

| | | |
|-----------|----------------------|-------------|
| ACCESSION | BX283642 | GI:28848096 |
| VERSION | BX283642.1 | |
| KEYWORDS | EST. | |
| SOURCE | Homo sapiens (human) | |
| ORGANISM | Homo sapiens | |

REFERENCE
AUTHORS
1 (bases 1 to 499)
Ebert, U., Hell, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Reich, H., Schaefer, D., and von P.

TITLE Human Unigeneset - RZPD3
JOURNAL Unpublished (2003)
COMMENT Contact: Tina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAG935K011707.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD (RZPDLIB No. 972)
<http://www.rzpd.de/cloncards/cgi-bin/showlib.pl.cgi?response=libNo=972> Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel.: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
PCMW-M13v, Primer sequence: CCGTGTAAACAGCGCCACT.

```

FEATURES
    source
        Location/Qualifiers
            1..499
                /organism="Homo sapiens"
                /mol_type="rRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE358K011707 ; IMAGE:4856568"
                /tissue_type="primary B-cells from tonsils (cell line)"
                /lab_host="DH10B (phage-resistant)"
                /clone_1b="NH1 MGC 48"
                /note="Organ: B-cells; Vector: pOT57; Site 1: XhoI;
                Site 2: EcoRI; cDNA made by oligo-dT priming.
                Directionally cloned into EcoRI/XhoI sites using the
                following 5' adaptor: GGCACTGAG(G). Size-selected >500bp
                for average insert size 1.8kb. Library constructed by Lin
                Hong in the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: this is a NH1_MGC library."

```

| | | | | | | | |
|-----------------------|--------|---|---------|------------|----|--------|-----|
| Query Match | 22.9% | Score | 498 | DB | 13 | Length | 499 |
| Best Local Similarity | 100.0% | Pred. No. | 2.9e-63 | | | | |
| Matches | 498 | Conservative | 0 | Mismatches | 0 | Indels | 0 |
| Gaps | | | | | | | |
| QY | 12 | GGTTGATGACTTACGAGAGCGCATTTGCTACTTTGGACCTTTGATATACCACTCCAGAAAT | 71 | | | | |
| DB | 2 | GGTTGATGACTTACGAGAGCGCATTTGCTACTTTGGACCTTTGATATACCACTCCAGAAAT | 61 | | | | |

| | | | |
|----|-----|--|-----|
| QY | 72 | GTCAACAAGATGAAATATCTGTGCGAAGACAGTGAATAAAAACAGTCCCTCAAGTCT | 131 |
| Db | 62 | GTCAACAACATGAAATATCTGTGCGAAGACAGTGAATAAAAACAGTCCCTCAAGTCT | 121 |
| QY | 132 | TCCTCGTTTTATCTTCAACTCTCACTTCTTGAAGTTACAGAAAAATTTATATA | 191 |
| Db | 122 | TCCTCGTTTTATCTTCAACTCTCACTTCTTGAAGTTACAGAAAAATTTATATA | 181 |
| QY | 192 | CGAATCTTTAAAAAGATCTATGTCTGAAAAATAGAGAAGAACACAGGTCTGGCCAGGA | 251 |
| Db | 182 | CGAATCTTTAAAAAGATCTATGTCTGAAAAATAGAGAAGAACACAGGTCTGGCCAGGA | 241 |
| QY | 252 | CGTGGTCAATTTGGTGCAGTTTGAATGCAACATGTGCCCTTACTGGAAATACAGAACT | 311 |
| Db | 242 | CGTGGTCAATTTGGTGCAGTTTGAATGCAACATGTGCCCTTACTGGAAATACAGAACT | 301 |
| QY | 312 | GCAGGACCTGGGAGCAATCCTTAAAGTCGAACGTTTTCTATGACTTTAGTAGGAGATGAG | 372 |
| Db | 302 | GCAGGACCTGGGAGCAATCCTTAAAGTCGAACGTTTTCTATGACTTTAGTAGGAGATGAG | 361 |
| QY | 372 | AGCAGAAGTAGATCCTTAAAAAGCATGATAGAGATCAATGTTTTTATATACAACATCC | 431 |
| Db | 362 | AGCAGAAGTAGATCCTTAAAAAGCATGATAGAGATCAATGTTTTTATATACAACATCC | 421 |
| QY | 432 | TTTATATTGATTCATTGAGTTAACAGTGGGTGTTAGTATAGATTTTCTATTTCTTTT | 491 |
| Db | 422 | TTTATATTGATTCATTGAGTTAACAGTGGGTGTTAGTATAGATTTTCTATTTCTTTT | 481 |
| QY | 492 | CCCTTGAAGTTACTTCC | 509 |
| Db | 482 | CCCTTGAAGTTACTTCC | 499 |

RESULT 14
 BG170756
 LOCUS
 DEFINITION
 BG170756 511 bp. mRNA linear. EST 06-FEB-2001
 6023287.F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4426960 5',
 mRNA sequence.
 ACCESSION
 BG170756
 KEYWORDS
 BG170756.1 GI:12677459
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 511)
 NIH-MGC <http://img.ncbi.nih.gov/>.
 JOURNAL
 National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-roman.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/BLMT at:
<http://image.llnl.gov>
 Plate: L14M10174 row: n column: 17
 High quality sequence stop: 511.
 Location/Qualifiers
 1..511

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4426960"
/tissue_type="hypernephroma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NH MGC 89"
/note="Organ: Kidney; Vector: pCMV-SPORT; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.3 kb. Library enriched for

```


full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

Query Match 22.7%; Score 493.6; DB 12; Length 511;
Best Local Similarity 99.0%; Pred. No. 1.2e-62;
Matches 507; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 696 ATAGACATGATGGACCTTCAAGCTACTTAATTAAGATCTTAAGATGGAGAG 755
DB 1 ATAGACATGATGGACCTTCAAGCTACTTAATTAAGATCTTAAGATGGAGAG 60
QY 756 GACTGTGAACAGACACCTTAATATGAGTGTGATGCTGAATAGCAATCTTCTGAAA 815
DB 61 GACTGTGAACAGACACCTTAATATGAGTGTGATGCTGAATAGCAATCTTCTGAAA 120
QY 816 CGCAAACTCTTTAAGAAAGTCCCTAATTTAGAAACCCCAAACTTCAATATCATPA 875
DB 121 CGCAAACTCTTTAAGAAAGTCCCTAATTTAGAAACCCCAAACTTCAATATCATPA 180
QY 876 TTAGCAAAATTTAGAAAGTGTGATGTTGGGAGAGAAAATCTATGCTCT 935
DB 181 TTAGCAAAATTTAGAAAGTGTGATGTTGGGAGAGAAAATCTATGCTCT 240
QY 936 CGTGGCTCTTCACTCAAGAAATGCAATGCAAGTTCGATCTTGTATG 995
DB 241 CGTGGCTCTTCACTCAAGAAATGCAATGCAAGTTCGATCTTGTATG 299
QY 996 TGTGATGCTTCTCCCAAGATATATTAATTAAGAGTGTGACAAACAGATGA 1055
DB 300 TGTGATGCTTCTCCCAAGATATATTAATTAAGAGTGTGACAAACAGATGA 359
QY 1056 TAAAGCTGGCAACCGTGGCAACGCTCATAGTTTGAAGTGTGGAGGTGAGAGGGA 1115
DB 360 TAAAGCTGGCAACCGTGGCAACGCTCATAGTTTGAAGTGTGGAGGTGAGAGGGA 419
QY 1116 GATGCTTGAACACAGGTGTTCAGAGCCAGCTGGGCAACATAAGATCTCTCT 1175
DB 420 GATGCTTGAACACAGGTGTTCAGAGCCAGCTGGGCAACATAAGATCTCTCT 479
QY 1176 CAAAAAAGAGAGAGAGAG 1207
DB 480 CAAAAAAGAGAGAGAGAG 511

RESULT 15
BGI70824 517 bp mRNA linear EST 06-FEB-2001
LOCUS BGI70824
DEFINITION 6022373F1 NIH_MGC_89 Homo sapiens CDNA clone IMAGE:4426955 5',
ACCESSION BGI70824
VERSION BGI70824.1 GI:12677527
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 517)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNA0174 row: n column: 16
High quality sequence stop: 517.
Location/Qualifiers

FEATURES

source

1. .517
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4426959"
/issue_type="hypertrophoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 89"
/note="Organ: kidney; Vector: pCMV-Sport6; Site: 1; NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

Query Match 21.7%; Score 471; DB 12; Length 517;
Best Local Similarity 97.8%; Pred. No. 2.4e-59;
Matches 488; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 708 TGTGACCTTCAAGCTACTTAATTAAGATCTTAAGATGGAGAGACTGTGAACA 767
DB 15 TGTGACCTTCAAGCTACTTAATTAAGATCTTAAGATGGAGAGACTGTGAACA 74
QY 768 GACACCTAATTAATGAGTGTGATGCTGAAGTACCAATCTTGTGAAGCAAACTCTT 827
DB 75 GACACCTAATTAATGAGTGTGATGCTGAAGTACCAATCTTGTGAAGCAAACTCTT 134
QY 828 TAAAGAGTCCCTAATTTAGAAACCCCAAACTTCAATATCATTAATTAAGCAAA 887
DB 135 TAAAGAGTCCCTAATTTAGAAACCCCAAACTTCAATATCATTAATTAAGCAAA 194
QY 888 TGAAGAGTGTGATGTTGGGAGAGAAAATCTATGCTCTGTGGCTCTT 947
DB 195 TGAAGAGTGTGATGTTGGGAGAGAAAATCTATGCTCTGTGGCTCTT 253
QY 948 CATCTGAAATGCAATGCAAGTTCAGAGTTGCTAATTTGATGTTGATGCTCT 1007
DB 254 CATCTGAAATGCAATGCAAGTTCAGAGTTGCTAATTTGATGTTGATGCTCT 313
QY 1008 CCAAGATATTAATTAATTAAGAGTGTGACAAACAGATGATTAAGCTGCGAA 1067
DB 314 CCAAGATATTAATTAATTAAGAGTGTGACAAACAGATGATTAAGCTGCGAA 373
QY 1068 CCGTGGCAACGCTCATAGTTCTAGCTGCTGGAGGTTGAGAGGAGATGCTTGA 1127
DB 374 CCGTGGCAACGCTCATAGTTCTAGCTGCTGGAGGTTGAGAGGAGATGCTTGA 433
QY 1128 CACAGGTGTCAAGGCGAGCTGGGCAACATTAAGATCTCTCTCAAAAAA 1187
DB 434 CACAGGTGTCAAGGCGAGCTGGGCAACATTAAGATCTCTCTCAAAAAA 493
QY 1188 AAAAAAAGAGAGAGAG 1206
DB 494 AAAAAAAGAGAGAGAG 512

Search completed: March 13, 2004, 00:39:08
Job time : 6920.32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2004, 17:20:29 ; Search time 66 seconds
(without alignments)
5019.783 Million cell updates/sec

Title: US-09-966-880A-7_COPY_80_676

Sequence: 1 atgcagacagcctctgatgtaa.....ttcgtacttggagcttga 597

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgm2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgm2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgm2_6/prodata/2/ina/5A_COMB.seq:*
4: /cgm2_6/prodata/2/ina/5B_COMB.seq:*
5: /cgm2_6/prodata/2/ina/PCTUS_COMB.seq:*
6: /cgm2_6/prodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--|
| 1 | 114.8 | 19.2 | 610 | 1 | US-08-816-241-2 Sequence 2, Appl |
| 2 | 114.8 | 19.2 | 610 | 3 | US-09-128-395-2 Sequence 2, Appl |
| 3 | 63.6 | 10.7 | 881 | 1 | US-08-687-895-2 Sequence 2, Appl |
| 4 | 63.6 | 10.7 | 891 | 2 | US-09-040-482-2 Sequence 2, Appl |
| 5 | 60 | 10.1 | 879 | 1 | US-08-158-682A-3 Sequence 3, Appl |
| 6 | 56.4 | 9.4 | 879 | 1 | US-08-158-682A-1 Sequence 1, Appl |
| 7 | 56.4 | 9.4 | 879 | 1 | US-08-015-203-1 Sequence 1, Appl |
| 8 | 38 | 6.4 | 1929 | 4 | US-09-380-420C-1 Sequence 1, Appl |
| 9 | 38 | 6.4 | 1929 | 4 | US-09-899-642A-1 Sequence 1, Appl |
| 10 | 37.8 | 6.3 | 922 | 3 | US-09-189-060B-73 Sequence 73, Appl |
| 11 | 36.6 | 6.1 | 4689 | 3 | US-09-105-537-34 Sequence 34, Appl |
| 12 | 36.6 | 6.1 | 36778 | 3 | US-09-105-537-5 Sequence 5, Appl |
| 13 | 36.6 | 6.1 | 38506 | 3 | US-09-320-878-19 Sequence 19, Appl |
| 14 | 36.6 | 6.1 | 38506 | 4 | US-09-141-908-1 Sequence 1, Appl |
| 15 | 36.6 | 6.1 | 38506 | 4 | US-09-657-440-19 Sequence 19, Appl |
| 16 | 36.4 | 6.1 | 624 | 4 | US-09-252-991A-10053 Sequence 10053, A |
| 17 | 36.4 | 6.1 | 1959 | 4 | US-09-252-991A-9783 Sequence 9783, Ap |
| 18 | 36.4 | 6.1 | 7218 | 1 | US-08-232-463-14 Sequence 14, Appl |
| 19 | 35.4 | 5.9 | 2190 | 4 | US-09-015-188-1 Sequence 1, Appl |
| 20 | 35.4 | 5.9 | 2584 | 4 | US-09-716-129-47 Sequence 47, Appl |
| 21 | 35 | 5.9 | 80246 | 3 | US-09-078-294-4 Sequence 4, Appl |
| 22 | 35 | 5.9 | 80595 | 3 | US-09-078-294-3 Sequence 3, Appl |
| 23 | 34.8 | 5.8 | 2505 | 4 | US-09-334-818A-1 Sequence 117, Appl |
| 24 | 34.6 | 5.8 | 1098 | 4 | US-09-170-496D-117 Sequence 117, Appl |
| 25 | 34.6 | 5.8 | 1098 | 4 | US-09-170-496D-225 Sequence 225, Ap |
| 26 | 34.6 | 5.8 | 1597 | 2 | US-08-724-974A-1 Sequence 1, Appl |
| 27 | 34.6 | 5.8 | 1697 | 4 | US-09-364-425B-26 Sequence 26, Appl |

| | | | | | |
|----|------|-----|---------|---|---------------------------------------|
| 28 | 34.6 | 5.8 | 2575 | 4 | US-09-716-129-25 Sequence 25, Appl |
| 29 | 34.4 | 5.8 | 2828 | 4 | US-09-016-434-1458 Sequence 1458, Ap |
| 30 | 34.4 | 5.8 | 2980 | 4 | US-09-266-225D-11 Sequence 11, Appl |
| 31 | 34.2 | 5.7 | 5159 | 2 | US-08-146-930-3 Sequence 3, Appl |
| 32 | 34.2 | 5.7 | 5159 | 3 | US-08-458-240-3 Sequence 3, Appl |
| 33 | 34.2 | 5.7 | 5159 | 5 | PCT-US93-03993-3 Sequence 3, Appl |
| 34 | 34.2 | 5.7 | 6359 | 4 | US-09-475-252-1 Sequence 1, Appl |
| 35 | 33.8 | 5.7 | 417 | 4 | US-09-252-991A-9297 Sequence 9297, Ap |
| 36 | 33.8 | 5.7 | 1215 | 4 | US-09-252-991A-9259 Sequence 9259, Ap |
| 37 | 33.8 | 5.7 | 1488 | 4 | US-09-091-609-1 Sequence 9270, Ap |
| 38 | 33.8 | 5.7 | 1602 | 3 | US-08-105-537-1 Sequence 1, Appl |
| 39 | 33.8 | 5.7 | 15872 | 4 | US-09-091-609-1 Sequence 1, Appl |
| 40 | 33.8 | 5.7 | 15872 | 4 | US-09-091-609-3 Sequence 3, Appl |
| 41 | 33.8 | 5.7 | 4403765 | 3 | US-09-103-840A-2 Sequence 2, Appl |
| 42 | 33.8 | 5.7 | 4411529 | 3 | US-09-103-840A-1 Sequence 1, Appl |
| 43 | 33.6 | 5.6 | 1683 | 4 | US-09-252-991A-8337 Sequence 8337, Ap |
| 44 | 33.6 | 5.6 | 1833 | 4 | US-09-252-991A-8371 Sequence 8371, Ap |
| 45 | 33.6 | 5.6 | 1833 | 4 | US-09-252-991A-8371 Sequence 8371, Ap |

ALIGNMENTS

RESULT 1
US-08-816-241-2
; Sequence 2, Application US/08816241
; Patent No. 5804185
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL RNA EDITING ENZYME
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/816,241
; FILING DATE: Filed Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-845-4166
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSTUT09
; CLONE: 1646833
; US-08-816-241-2

Query Match 19.2%; Score 114.8; DB 1; Length 610;
Best Local Similarity 53.9%; Pred. No. 1,76-25;
Matches 289; Conservative 0; Mismatches 223; Indels 15; Gaps 2;

QY 16 ATGACCGGAGGAGAGTTCTTTTACCAATTCAAAAATGTCGGCTAGGGCTAGGGCT 75
DB 34 ATGTATCCAGGACATTTCTACTTCAATTTAAACCTATGGAAAGCCACATCGGAC 93
QY 76 GAGACCTACCTGTGTACAGTAGTGAAGAGCGGTGACATGTCTACATCTTTTCACTGGAC 135
DB 94 GAAACTTGGCTGTGCTTCAACCGTGAAGAGATTAAGAGCGCGCTCAGTTGTCTCTGGAAAG 153
QY 136 TTGTGTTATCTTGGCAA-----TAAAGACGGCTGCCAGTGGAAATGCTCTTC 183
DB 154 ACGGCGGTCTTCCGAAACAGGTGATCTGTGAGACCATGTCTATGCAAGAAAGTGTCTTC 213
QY 184 CTCCGCTACATCTCGAGCTGGAGCCTAGACCCCTGGCGGCTGCTACCGGCTCAGCTGTTC 243
DB 214 CTCTCTGTGCTTGGCAGCAGCATATCTCTCTTAACACAAAGTACAGGTCTCAGTGTAC 273
QY 244 ACCTCTGAGCCCTCTGCTACGACTGTGCCGACATGTGCGGACTTTCTGCGAGGAGAC 303
DB 274 ACATCTTGGAGCCCTTGGCCAGCTGTGAGGGAGTGGCCGAGTTCTGTGCGCAGGAC 333
QY 304 CCCACCTCAGTCTGAGATCTTACCGCGCGCTCTACTTGTGAGAGCCGCAAGCT 363
DB 334 AGCAAGTGAATCTCACTCACTTCAACCGCGCTCTACTTCCAGTATC---CATGT 390
QY 364 GAGCCGAGGGGCTGGCGGCTGCAACCGCGCGGCTGCAATAGCCATCATGACTTC 423
DB 391 TACCAAGAGGGGCTCCGACGCTGAGTCAAGAGGGGTGCTGTGAGTCACTGAGACTAT 450
QY 424 AAGATTAATTTTACTGCTGGAATACTTTTGTAGAAACATGAAAGAACTTTCAAGCC 483
DB 451 GAAGATTTTAATATTTGTGGGAAACTTTGTGTACATGATATAGACCATTCAGCTT 510
QY 484 TGGAGAGGGCTGCATGAAATTCAGTCTGTCTCCAGACAGCTTCGCGCATCTT 539
DB 511 TGGAGGATTAAACCACTTTCGACTTGTGAAAGAGGCTACGGAGAGTCT 566

RESULT 2
US-09-128-395-2
Sequence 2, Application US/09128395
Patent No. 6087108
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL RNA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/128,395
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/816,241
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PP-0239 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 610 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSTUT09
CLONE: 1646833
US-09-128-395-2

Query Match
Best Local Similarity 53.8%; Pred. No. 1.7e-25;
Matches 289; Conservative 0; Mismatches 232; Indels 15; Gaps 2;

QY 16 ATGACCGGAGGAGAGTTCTTTTACCAATTCAAAAATGTCGGCTAGGGCTAGGGCT 75
DB 34 ATGTATCCAGGACATTTCTACTTCAATTTAAACCTATGGAAAGCCACATCGGAC 93
QY 76 GAGACCTACCTGTGTACAGTAGTGAAGAGCGGTGACATGTCTACATCTTTTCACTGGAC 135
DB 94 GAAACTTGGCTGTGCTTCAACCGTGAAGAGATTAAGAGCGCGCTCAGTTGTCTCTGGAAAG 153
QY 136 TTGTGTTATCTTGGCAA-----TAAAGACGGCTGCCAGTGGAAATGCTCTTC 183
DB 154 ACGGCGGTCTTCCGAAACAGGTGATCTGTGAGACCATGTCTATGCAAGAAAGTGTCTTC 213
QY 184 CTCCGCTACATCTCGAGCTGGAGCCTAGACCCCTGGCGGCTGCTACCGGCTCAGCTGTTC 243
DB 214 CTCTCTGTGCTTGGCAGCAGCATATCTCTCTTAACACAAAGTACAGGTCTCAGTGTAC 273
QY 244 ACCTCTGAGCCCTCTGCTACGACTGTGCCGACATGTGCGGACTTTCTGCGAGGAGAC 303
DB 274 ACATCTTGGAGCCCTTGGCCAGCTGTGAGGGAGTGGCCGAGTTCTGTGCGCAGGAC 333
QY 304 CCCACCTCAGTCTGAGATCTTCAACCGCGCGCTCTACTTGTGAGAGCCGCAAGCT 363
DB 334 AGCAAGTGAATCTCACTCACTTCAACCGCGCTCTACTTCCAGTATC---CATGT 390
QY 364 GAGCCGAGGGGCTGGCGGCTGCAACCGCGCGGCTGCAATAGCCATCATGACTTC 423
DB 391 TACCAAGAGGGGCTCCGACGCTGAGTCAAGAGGGGTGCTGTGAGTCACTGAGACTAT 450
QY 424 AAGATTAATTTTACTGCTGGAATACTTTTGTAGAAACATGAAAGAACTTTCAAGCC 483
DB 451 GAAGATTTTAATATTTGTGGGAAACTTTGTGTACATGATATAGACCATTCAGCTT 510
QY 484 TGGAGAGGGCTGCATGAAATTCAGTCTGTCTCCAGACAGCTTCGCGCATCTT 539
DB 511 TGGAGGATTAAACCACTTTCGACTTGTGAAAGAGGCTACGGAGAGTCT 566

RESULT 3
US-08-687-895-2
Sequence 2, Application US/08687895
Patent No. 5747319
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Hawkins, Phillip R.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: A NOVEL HUMAN MENA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5


```

; Patent No. 5434058
; GENERAL INFORMATION:
; APPLICANT: Davidson, Nicholas O.
; TITLE OF INVENTION: Apolipoprotein B RNA Editing Protein:
; NUMBER OF SEQUENCES: 18
; TITLE OF INVENTION: Composition and Method
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: 321 No. 5434058th Clark Street, Suite 800
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60610
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/158,682A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooley, Ronald B.
; REGISTRATION NUMBER: 27,187
; REFERENCE/DOCKET NUMBER: ARCD:085
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 744-0090
; TELEFAX: (312) 245-4961
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 23..730
; US-08-158-682A-3

Query Match          10.1%; Score 60; DB 1; Length 879;
Best Local Similarity 54.1%; Pred. No. 1.4e-08;
Matches 145; Conservative 0; Mismatches 120; Indels 3; Gaps 1;

QY 196 TCGGCTGGGACCTAGAGACCTGGCGCTGCTACCGGCTCACTGTTCACTCTCTGGAGC 255
DB 236 TCAAGAAAGATTTTCAACCCATCATAGCTGCACCATCACTGCTCTGTCTGTGAGT 295
QY 256 CCTGCTACGACTGTGCCGACATGTGACCGACTTTTGGCAGGGAACCCCAACTGAGT 315
DB 296 CCTGCTGGGAAATGCTCCAGAGTATAGAGATTTTGAAGTGGGACCTGTGTGACT 355
QY 316 CTGAGATCTTACCGGCGGCTCTACTCTGTAGAGACCGGAAGGTTGAGCCGAGGG 375
DB 356 CTAGTATCTAGTACGCTGCGCTTTTGGCAGCATGATCAACAAATCGGC--AAGGT 412
QY 376 CTGGGGGGGCTGCACCGCGCGGGGGTGAATAGCCATGAGCTTCAAGATATATTT 435
DB 413 CTCAGGAGCTTTGTTAAAGTGAAGTAACTATTAAGATTAAGAGCATCAGAGTATTA 472
QY 436 TACTGCTGGAATACTTTTGTAGAAACC 463
DB 473 CACTGCTGAGAGATTTTGTCACTACC 500

```

```

; TITLE OF INVENTION: Composition and Method
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: 321 No. 5434058th Clark Street, Suite 800
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60610
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/158,682A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooley, Ronald B.
; REGISTRATION NUMBER: 27,187
; REFERENCE/DOCKET NUMBER: ARCD:085
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 744-0090
; TELEFAX: (312) 245-4961
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-158-682A-1

Query Match          9.4%; Score 56.4; DB 1; Length 879;
Best Local Similarity 53.4%; Pred. No. 1.7e-07;
Matches 142; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

QY 197 CGAGTGGGACCTAGAGACCTGGCGGCTGCTACCGGCTCACTGTTCACTCTCTGGAGC 256
DB 255 CAGAAAGATACTTTTGTCCAAACACAGATGCTCATTAACCTGCTCTCTCTGGAGTC 314
QY 257 CCTGCTACGACTGTGCCGACATGTGACCGACTTTTGGCAGGGAACCCCAACTGAGT 316
DB 315 CCTGCTGGGAAATGCTCCAGAGTATAGAGATTTTGAAGTGGGACCTGTGTGACT 374
QY 317 TGAGGATCTTACCGGCGGCTCTACTCTGTGAGGACCGGCAAGCTGAGCCGAGGG 376
DB 375 TGTATTATTAATAGACAGGCTTTATACACAGAGATCTCGAAATCGGC--AAGAG 431
QY 377 TGGCGGCGCTGCACCGCGCGGGGTGAATAGCCATGATCACTTCAAGATATATTT 436
DB 432 TCAGGACACTATTATAGACGCGGTGTACTATCCAGATCATAGCAGAGAGAGTCTGCT 491
QY 437 ACTGCTGGAATCTTTGTAGAAACC 462
DB 492 ACTGCTGAGAGATTTTGTCACTAC 517

```

```

RESULT 6
US-08-158-682A-1
; Sequence 1, Application US/08158682A
; Patent No. 5434058
; GENERAL INFORMATION:
; APPLICANT: Davidson, Nicholas O.
; TITLE OF INVENTION: Apolipoprotein B RNA Editing Protein:
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: 321 No. 5550034th Clark Street, Suite 800

```

```

1 MEDIUM TYPE: Floppy disk
2 COMPUTER: IBM PC compatible
3 OPERATING SYSTEM: PC-DOS/MS-DOS
4 SOFTWARE: Patent in Release #1.0, Version #1.25
5
6 CURRENT APPLICATION DATA:
7 APPLICATION NUMBER: US/09/380,420C
8 FILING DATE: 12-NO. 6300544-1999
9
10 CLASSIFICATION: <Unknown>
11
12 ATTORNEY/AGENT INFORMATION:
13 NAME: Meigs, J. Timothy
14 REGISTRATION NUMBER: 38,241
15 REFERENCE/DOCKET NUMBER: S-21251A
16 TELECOMMUNICATION INFORMATION:
17 TELEPHONE: 919-541-8587
18
19 INFORMATION FOR SEQ ID NO: 1:
20
21 SEQUENCE CHARACTERISTICS:
22
23 LENGTH: 1929 base pairs
24 TYPE: nucleic acid
25 STRANDEDNESS: double
26 TOPOLOGY: linear
27
28 MOLECULE TYPE: CDNA
29 IMMEDIATE SOURCE:
30 CLONE: P450ox
31
32 FEATURE:
33 NAME/KEY: CDS
34 LOCATION: 81..1673
35
36 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
37
38 JS-09-380-420C-1

```

| | | | | |
|-----------------------|--------|-----------------|-------|----------------------------------|
| Query Match | 6.4%; | Score 38; | DB 4; | Length 1929; |
| Best Local Similarity | 53.3%; | Pred. No. 0.11; | | |
| Matches | 80; | Conservative | 0; | Mismatches 70; Indels 0; Gaps 0; |

| | | | |
|----|-----|---|-----|
| Qy | 236 | CCTGTTAACTCCTCGAGGCCCTGCTACACATGTGCCGACATGTGCCGACCTTTTCGC | 295 |
| | | | |
| Db | 793 | CCAGCTTTCCGCCGAGGACTTCTTCCCAACGCCGCCGCGCCTCGCGAGCCGCTTCT | 852 |
| | | | |
| Qy | 296 | GAGGAAACCCCACTCAGTCTGAGATCTTCAACCGCGGCTCTACTTCTGTGAGACC | 355 |
| | | | |
| Db | 853 | CGGCTTCTCGCGCGCGCGAGGAGCATCTTCAACAGACTCGACGCTTCTTCGAGAGG | 912 |
| | | | |
| Qy | 356 | GCAAGCTGAGCCCGAGGCGCTGCGCGGC | 385 |
| | | | |
| Db | 913 | TCATGACCAAGACATGGAACCGCGCGGC | 942 |
| | | | |

| | |
|--|--|
| RESULT 9 | |
| US-09-899-642A-1 | |
| / Sequence 1, Application US/09899642A | |
| / Patent No. 6649814 | |
| / GENERAL INFORMATION: | |
| / | |
| APPLICANT: Halkier, Barbara | |

1 Moller, Birger
 2 TITLE OF INVENTION: Cytochrome P450 Monooxygenases
 3 NUMBER OF SEQUENCES: 23
 4 CORRESPONDENCE ADDRESS:
 5 ADDRESSEE: Syngenta Patent Dept.
 6 STREET: 3054 Cornwallis Road
 7 CITY: RTP
 8 STATE: NC
 9 COUNTRY: USA
 10 ZIP: 27709
 11
 12 COMPUTER READABLE FORM:
 13 MEDIUM TYPE: floppy disk
 14 COMPUTER: IBM PC compatible
 15 OPERATING SYSTEM: PC-DOS/MS-DOS
 16 SOFTWARE: Patentin Release #1.0, Version #1.25
 17
 18 CURRENT APPLICATION DATA:
 19 APPLICATION NUMBER: US/09/899,642A
 20 FILING DATE: 05-Jul-2001
 21 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/380,420
FILING DATE: 12-NO. 6649814-1999
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-21251A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1929 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: P4500x
FEATURE:
NAME/KEY: CDS
LOCATION: 81..1673
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-899-642A-1

Query Match 6.4%; Score 38; DB 4; Length 1929;
Best Local Similarity 53.3%; Pred. No. 0.11;
Matches 80; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 236 CCGGTTCACCTCCGTGAGAGCCCTGTACAGCTGAGCCGACATGAGCCGACCTTTTCGC 295
DB 793 CCAAGCTTCCTCCGCGAGACTTCTTCCCAAGCCGCGCCGCTCGCCGACCGCTCT 852
QY 296 GAGGGAACCCCACTCACTGAGATCTTCAACCGCGCCCTCTACTTCTGTGAGACC 355
DB 853 CGGCTTCCTCCGCGCCCGCGAGCGCATCTTCAAGAGCTGACGCTTCTTCCGAGAGG 912
QY 356 GCAAGGCTGAGCCCGAGGCGCTGCGCGGC 385
DB 913 TCATGACCAAGCATGAGACCGCGCGCGCC 942

RESULT 10

US-09-189-060B-73
Sequence 73, Application US/09189060B

PATENT NO. 6270968
GENERAL INFORMATION:
APPLICANT: Dalboge, Henrik
APPLICANT: Sandel, Thomas
APPLICANT: Kauppinen, Markus
APPLICANT: Borge, Diderichsen
TITLE OF INVENTION: Method Of Providing No. 6270968e1 DNA Sequences
FILE REFERENCE: 4772.204-US
CURRENT APPLICATION NUMBER: US/09/189,060B
CURRENT FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: PCT/DK97/00216
PRIOR FILING DATE: 1997-05-12
NUMBER OF SEQ ID NOS: 74
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 73
LENGTH: 922
TYPE: DNA
ORGANISM: Hybrid
FEATURE:
NAME/KEY: CDS
LOCATION: (2)...(922)
US-09-189-060B-73

Query Match 6.3%; Score 37.8; DB 3; Length 922;
Best Local Similarity 49.3%; Pred. No. 0.091;
Matches 99; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 176 TGCTCTTCCTCCGCTACATCTCGAGCTGAGACCTAGACCTTGCGCGCTGCTACCGGCTCA 235

DB 378 TGATGCTCCAGTCCACCAACAGGCGCGGACCTGTCCGACCAACCTTGCAGCTGCTCA 437
QY 236 CCGGTTCACCTCCGTGAGAGCCCTGTACAGCTGAGCCGACATGAGCCGACCTTCTGC 295
DB 438 TCCCGCGCGCGCGCTGCGCATTTGACGCTGACCTTCCAGTACGCGCCAGCCCTTCG 497
QY 296 GAGGGAACCCCACTCACTGAGATCTTCAACCGCGCCCTCTACTTCTGTGAGACC 355
DB 498 CCGCGCGCGCGCTGCGCATTTGACGCTCCCGCGCGAGTGCAGATGCCGAGGCCA 557
QY 356 GCAAGGCTGAGCCCGAGGCGGC 376
DB 558 TCAAGCGCGCTGCCAGTGC 578

RESULT 11

US-09-105-537-34
Sequence 34, Application US/09105537A

PATENT NO. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/105,537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 34
LENGTH: 4689
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-105-537-34

Query Match 6.1%; Score 36.6; DB 3; Length 4689;
Best Local Similarity 46.9%; Pred. No. 0.48;
Matches 114; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 156 GAAGCGCTGCGACGTGAGATTGCTCTCCGCTACATCTGAGACTGAGACCTAGACCC 215
DB 3402 GATGCGCCACACGAGAGCCGAACACTCTCTCTGTCAAGCGAGCGGAGACACCC 3461
QY 216 TGCGCGCTGCTACCGCTCACTGATTCACCTTCGAGCCCTGTACGATGTGCGCG 275
DB 3462 CGAGGCCACCACTACCGCGCGGACGATCGGCGCGCGCTACCATCGCCCG 3521
QY 276 ACATGTGCGGACTTTTGTGAGAGGAACCCCACTTCACTGATGATCTTCAACCGCG 335
DB 3522 CTGCGAGCTGCGGACCCCGACGCGCATGCGGACCTCTCTGAGCGCATCCCGCGAGAC 3581
QY 336 CCGTACTTCTGTGAGAGCGGAGGCTGAGCCCGGAGGCGCTCGGCGGCTGACCGCG 395
DB 3582 GCCCTTACCGCGCTCTTCAACCGCGCGGACCGGCGGAGATCGCTGACGCTCAC 3641
QY 396 CGG 398
DB 3642 CGG 3644

RESULT 12

US-09-105-537-5
Sequence 5, Application US/09105537A

PATENT NO. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/105,537A

;; CURRENT FILING DATE: 1998-06-26
;; NUMBER OF SEQ ID NOS: 43
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 5
;; LENGTH: 36778
;; TYPE: DNA
;; ORGANISM: Streptomyces venezuelae
US-09-105-537-5

Query Match 6.1%; Score 36.6; DB 3; Length 36778;
Best Local Similarity 46.9%; Pred. No. 1.4;
Matches 114; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 156 GAACGGCTGCACGTGAGATTGCTCTCCGCTACATCTCGAGCTGGAGCTAGACCC 215
DB 30392 GATGGCCACCAACCGAGCCGAAACCTCTCTCTCTGTCAGCCGACGCGCAACAGCCCC 30451
QY 216 TGGCCGCTGTACCGCGTCACTGTGTACCTCTGTAGAGCCCTGTACGACTGTGCGG 275
DB 30452 CGGAGCCACCACTACCGCGGCACTACCGCATCGGGCGCCGCTACACATCGCCG 30511
QY 276 ACATGTGCGGACTTCTTGTGAGAGAAACCCCACTTCACTGATGAGATCTTCAACGCGG 335
DB 30512 CTGGAGCGTGCAGACCCCGACCGCATGCGACCTCTCTGACGCCATCCCGCGAGAC 30571
QY 336 CCTTACTTCTGTGAGAGACCGCAAGGTGAGCCCGAGGGGCTGCGCGGTGACCGCGG 395
DB 30572 GCGCCCTACCGCGCTGTCTTCCACACCGCGGCGGACCGGGCGGATCGCTGACGCTAC 30631
QY 396 CGG 398
DB 30632 CGG 30634

RESULT 13

US-09-320-878-19
; Sequence 19; Application US/09320878A
; Patent No. 6117659

;; GENERAL INFORMATION:
;; APPLICANT: ASHLEY, Gary
;; APPLICANT: BETLACH, Melanie C.
;; APPLICANT: BETLACH, Mary C.
;; APPLICANT: MCDANIEL, Robert

;; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
;; FILE REFERENCE: 30622002120
;; CURRENT APPLICATION NUMBER: US/09/320, 878A

;; EARLIER FILING DATE: 1998-05-27
;; EARLIER APPLICATION NUMBER: CIP OF 09/141, 908
;; EARLIER FILING DATE: 1998-08-28
;; EARLIER APPLICATION NUMBER: CIP OF 09/073, 538
;; EARLIER FILING DATE: 1998-05-06
;; EARLIER APPLICATION NUMBER: CIP OF 08/846, 247
;; EARLIER FILING DATE: 1997-04-30
;; EARLIER APPLICATION NUMBER: 60/119, 139
;; EARLIER FILING DATE: 1999-02-08
;; EARLIER APPLICATION NUMBER: 60/100, 880
;; EARLIER FILING DATE: 1998-09-22
;; EARLIER APPLICATION NUMBER: 60/087, 080
;; EARLIER FILING DATE: 1998-05-28
;; NUMBER OF SEQ ID NOS: 34
;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 19
;; LENGTH: 38506
;; TYPE: DNA
;; ORGANISM: Streptomyces venezuelae
US-09-320-878-19

Query Match 6.1%; Score 36.6; DB 3; Length 38506;
Best Local Similarity 46.9%; Pred. No. 1.4; Indels 129; Indels 0; Gaps 0;
Matches 114; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 156 GAACGGCTGCACGTGAGATTGCTCTCTCCGCTACATCTCGAGCTGGAGCTAGACCC 215

DB 28534 GATGGCCACCAACCGAGCCGAAACCTCTCTCTCTGTCAGCCGACGCGCAACAGCCCC 28593

QY 216 TGGCCGCTGTACCGCGTCACTGTGTACCTCTGTAGAGCCCTGTACGACTGTGCGG 275

DB 28594 CGGAGCCACCACTACCGCGGCACTACCGCATCGGGCGCCGCTACACATCGCGG 28653

QY 276 ACATGTGCGGACTTCTTGTGAGAGAAACCCCACTGATGAGATCTTCAACCGCGG 335

DB 28654 CTGGAGCGTGCAGACCCCGACCGCATGCGACCTCTGACGCCATCCCGCGAGAC 28713

QY 336 CCTTACTTCTGTGAGAGACCGCAAGGTGAGCCCGAGGGGCTGCGCGGTGACCGCGG 395

DB 28714 GCGCCCTACCGCGCTGTCTTCCACACCGCGGCGGACCGGGCGGATCGCTGACGCTAC 28773

QY 396 CGG 398
DB 28774 CGG 28776

RESULT 14

US-09-141-908-1

;; Sequence 1, Application US/09141908
;; Patent No. 6503741

;; GENERAL INFORMATION:

;; APPLICANT: ASHLEY, Gary
;; APPLICANT: BETLACH, Melanie C.
;; APPLICANT: BETLACH, Mary
;; APPLICANT: MCDANIEL, Robert

;; APPLICANT: TANG, Li

;; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a
;; FILE OF INVENTION: Modular PKs Gene Cluster as Scaffold

;; FILE REFERENCE: 30622002100
;; CURRENT APPLICATION NUMBER: US/09/141, 908

;; EARLIER FILING DATE: 1998-08-28
;; EARLIER APPLICATION NUMBER: CIP OF 09/073, 538

;; EARLIER FILING DATE: 1998-05-06
;; EARLIER APPLICATION NUMBER: CIP OF 08/846, 247

;; EARLIER FILING DATE: 1997-04-30
;; EARLIER APPLICATION NUMBER: PROV. 60/076, 919

;; EARLIER FILING DATE: 1998-03-05
;; EARLIER APPLICATION NUMBER: PROV. 60/087, 080

;; EARLIER FILING DATE: 1998-05-28
;; NUMBER OF SEQ ID NOS: 31

;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 1

;; LENGTH: 38506
;; TYPE: DNA

;; ORGANISM: Streptomyces venezuelae
US-09-141-908-1

Query Match 6.1%; Score 36.6; DB 4; Length 38506;
Best Local Similarity 46.9%; Pred. No. 1.4;
Matches 114; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 156 GAACGGCTGCACGTGAGATTGCTCTCTCCGCTACATCTCGAGCTGGAGCTAGACCC 215

DB 28534 GATGGCCACCAACCGAGCCGAAACCTCTCTCTGTCAGCCGACGCGCAACAGCCCC 28593

QY 216 TGGCCGCTGTACCGCGTCACTGTGTACCTCTGTAGAGCCCTGTACGACTGTGCGG 275

DB 28594 CGGAGCCACCACTACCGCGGCACTACCGCATCGGGCGCCGCTACACATCGCGG 28653

QY 276 ACATGTGCGGACTTCTTGTGAGAGAAACCCCACTGATGAGATCTTCAACCGCGG 335

DB 28654 CTGGAGCGTGCAGACCCCGACCGCATGCGACCTCTTGAAGCCATCCCGCGAGAC 28713

QY 336 CCTTACTTCTGTGAGAGACCGCAAGGTGAGCCCGAGGGGCTGCGCGGTGACCGCGG 395

DB 28714 GCGCCCTACCGCGCTGTCTTCCACACCGCGGCGGACCGGGCGGATCGCTGACGCTAC 28773

QY 396 CGG 398
DB 28774 CGG 28776

Db 28774 CGG 28776

RESULT 15
US-09-657-440-19
; Sequence 19; Application US/09657440
; Patent No. 6509455
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/657,440
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: CIP OF 09/141,908
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-657-440-19

Query Match 6.1%; Score 36.6; DB 4; Length 38506;
Best Local Similarly 46.9%; Pred. No. 1.4;
Matches 114; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

| | | | |
|----|-------|--|-------|
| QY | 156 | GAACGGCTGCACGTGGAATTGCTCTCTCCCGCTACATCGAGCTGGAGCTAGAGCC | 215 |
| Db | 28534 | GATGGCCACCGAGAGCCGGAACACCTCTCTCTGTCAGCCGACGCGGACAGACCC | 28593 |
| QY | 216 | TGCGCGCTGCTACCGCGCTCACTGATTCACCTCCTGGAGCCCTGCTACGACTGTGCCG | 275 |
| Db | 28594 | CGGAGCCACCACTCACTCAGCCGCACTACCGCATCGGCGCCGCGTCACTGCGCCG | 28653 |
| QY | 276 | ACATGTGCGCACTTCTTGCGAGGAAACCCCACTCACTGAGATCTTCAACCGCGC | 335 |
| Db | 28654 | CTGCGAGTGCAGCCGACCCGACGCGCATGCGCACCTCTCGACGCACTCCCGCGAGAC | 28713 |
| QY | 336 | CCTTACTTCTGTGAGAGACCGCAAGGCTGAGCCCGAGGGGCTGGCGGGCTGCACCGCG | 395 |
| Db | 28714 | GCCCCCTACCGCCGTCGTCACACCGCGGCGCACCGGCGGCGGATCCGCTGAGCTCAC | 28773 |
| QY | 396 | CGG 398 | |
| Db | 28774 | CGG 28776 | |

Search completed: March 12, 2004, 09:40:12
Job time : 74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2004, 09:15:07 ; Search time 248 Seconds
(without alignments)
8862.556 Million cell updates/sec

Title: US-09-966-880A-7_COPY_80_676
Sequence: 1 atggacacagctcttgatgaa.....ttcgracttgggacttga 597

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2432557 seqs, 1840798864 residues

Total number of hits satisfying Chosen parameters: 4865114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 597 | 100.0 | 2818 | 9 | US-09-966-880A-7 |
| 2 | 475.4 | 79.6 | 2440 | 9 | US-09-966-880A-1 |
| 3 | 273.4 | 45.8 | 6564 | 9 | US-09-966-880A-10 |
| 4 | 273.4 | 45.8 | 11204 | 9 | US-09-966-880A-35 |
| 5 | 271 | 45.4 | 271 | 9 | US-09-966-880A-13 |
| 6 | 148 | 24.8 | 148 | 9 | US-09-966-880A-12 |
| 7 | 118.4 | 19.8 | 1155 | 15 | US-10-460-923-1 |
| 8 | 118.4 | 19.8 | 1534 | 9 | US-09-729-674-173 |
| 9 | 114.8 | 19.2 | 987 | 9 | US-09-925-300-699 |
| 10 | 114.8 | 19.2 | 1143 | 13 | US-10-098-841-268 |
| 11 | 114.4 | 19.2 | 116 | 9 | US-09-966-880A-14 |
| 12 | 110.4 | 18.5 | 1348 | 9 | US-09-880-107-3282 |
| 13 | 107.2 | 18.0 | 823 | 14 | US-10-247-671-62 |
| 14 | 107.2 | 18.0 | 1008 | 15 | US-10-131-827-8890 |
| 15 | 104.4 | 17.5 | 2151 | 9 | US-09-822-830A-359 |

| | | | | | |
|----|------|------|--------|----|----------------------|
| 16 | 90.2 | 15.1 | 2913 | 15 | US-10-027-632-112023 |
| 17 | 90.2 | 15.1 | 2913 | 15 | US-10-027-632-112024 |
| 18 | 90.2 | 15.1 | 201143 | 12 | US-10-240-425-1039 |
| 19 | 78.4 | 13.1 | 371 | 10 | US-09-918-995-30237 |
| 20 | 71.6 | 12.0 | 1404 | 14 | US-10-198-846-13472 |
| 21 | 71.6 | 12.0 | 2773 | 15 | US-10-104-047-1759 |
| 22 | 70.6 | 11.8 | 456 | 14 | US-10-029-386-25187 |
| 23 | 70.2 | 11.8 | 390 | 9 | US-09-796-692-9051 |
| 24 | 70.2 | 11.8 | 390 | 14 | US-10-040-862-9051 |
| 25 | 70.2 | 11.8 | 390 | 15 | US-10-057-475B-9051 |
| 26 | 70.2 | 11.8 | 390 | 15 | US-10-154-884B-9051 |
| 27 | 67 | 11.2 | 572 | 9 | US-09-864-761-9416 |
| 28 | 66.6 | 11.2 | 328 | 9 | US-09-796-692-5737 |
| 29 | 66.6 | 11.2 | 328 | 14 | US-10-040-862-5737 |
| 30 | 66.6 | 11.2 | 328 | 15 | US-10-057-475B-5737 |
| 31 | 66.6 | 11.2 | 328 | 15 | US-10-154-884B-5737 |
| 32 | 63 | 10.6 | 476 | 10 | US-09-918-995-32185 |
| 33 | 60 | 10.1 | 879 | 14 | US-10-157-031-13 |
| 34 | 57 | 9.5 | 445 | 15 | US-10-378-029-25 |
| 35 | 54 | 9.0 | 2172 | 9 | US-09-966-880A-15 |
| 36 | 53.2 | 8.9 | 821 | 15 | US-10-388-934-189 |
| 37 | 50.2 | 8.4 | 674 | 14 | US-10-257-826A-150 |
| 38 | 49.4 | 8.3 | 556 | 9 | US-09-796-692-5061 |
| 39 | 49.4 | 8.3 | 556 | 14 | US-10-040-862-5061 |
| 40 | 49.4 | 8.3 | 556 | 15 | US-10-057-475B-5061 |
| 41 | 49.4 | 8.3 | 556 | 15 | US-10-154-884B-5061 |
| 42 | 48 | 8.0 | 553 | 9 | US-09-796-692-4504 |
| 43 | 48 | 8.0 | 553 | 14 | US-10-040-862-4504 |
| 44 | 48 | 8.0 | 553 | 15 | US-10-057-475B-4504 |
| 45 | 48 | 8.0 | 553 | 15 | US-10-154-884B-4504 |

ALIGNMENTS

RESULT 1
US-09-966-880A-7
Sequence 7, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Taku
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966, 880A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 2818
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (80)...(673)
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)...(79)
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (677)...(2818)
US-09-966-880A-7
Query Match 100.0%; Score 597; DB 9; Length 2818;
Best Local Similarity 100.0%; Pred. No. 1.4e-187;

Sequence 112023,
Sequence 112024,
Sequence 1099, Ap
Sequence 30237, A
Sequence 13472, A
Sequence 1759, A
Sequence 25187, A
Sequence 9051, Ap
Sequence 9051, Ap
Sequence 9051, Ap
Sequence 9051, Ap
Sequence 9416, Ap
Sequence 5737, Ap
Sequence 5737, Ap
Sequence 5737, Ap
Sequence 5737, Ap
Sequence 32185, A
Sequence 13, Appl
Sequence 25, Appl
Sequence 15, Appl
Sequence 189, Appl
Sequence 150, Ap
Sequence 5061, Ap
Sequence 5061, Ap
Sequence 5061, Ap
Sequence 5061, Ap
Sequence 4504, Ap
Sequence 4504, Ap
Sequence 4504, Ap
Sequence 4504, Ap

Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGACAGCCTCTGATGAAACCGAGAGATTCTTTACCAATTCAAAATATGCCCTGG 60
Db 80 ATGACAGCCTCTGATGAAACCGAGAGATTCTTTACCAATTCAAAATATGCCCTGG 139
QY 61 GCTAAGGATCGCGTGAAGACCTACCTGTGTAAGTAAGAGCGGTGACAGTGTACA 120
Db 140 GCTAAGGATCGCGTGAAGACCTACCTGTGTAAGTAAGAGCGGTGACAGTGTACA 199
QY 121 TCTTTTCTACTGACTTTGTGTATCTTCCGATTAAGAAAGCGTCCACCTGGAATTGCTC 180
Db 200 TCTTTTCTACTGACTTTGTGTATCTTCCGATTAAGAAAGCGTCCACCTGGAATTGCTC 259
QY 181 TTCTCCGCTACATCTCGGACTGGGACCTAGACCTGGCGCGCTGCTACCGCGTCACTGG 240
Db 260 TTCTCCGCTACATCTCGGACTGGGACCTAGACCTGGCGCGCTGCTACCGCGTCACTGG 319
QY 241 TTCACTCTGAGACCCCTGCTACGACTGTGCCGACATGTGGCCGACATTTCTGCGAGGG 300
Db 320 TTCACTCTGAGACCCCTGCTACGACTGTGCCGACATGTGGCCGACATTTCTGCGAGGG 379
QY 301 AACCCCAACTCACTGCTGAGGATCTTCAACCGCGCGCTTACTTCTGTGAGAGACGCAAG 360
Db 380 AACCCCAACTCACTGCTGAGGATCTTCAACCGCGCGCTTACTTCTGTGAGAGACGCAAG 439
QY 361 GCTGAGCCGAGGGGCTGCGCGGCTGCGACCGCGCGGGGTGCAAAATAGCATATGACC 420
Db 440 GCTGAGCCGAGGGGCTGCGCGGCTGCGACCGCGCGGGGTGCAAAATAGCATATGACC 499
QY 421 TTCAAGATATTTTATCTGCTGAGAACTTTGTGAGAAAACCATGAAGAACTTTCAAA 480
Db 500 TTCAAGATATTTTATCTGCTGAGAACTTTGTGAGAAAACCATGAAGAACTTTCAAA 559
QY 481 GCCTGGAGAGGCTGCATGAAATTCAGTTCTCTCCAGACAGCTTGGCGCATCTT 540
Db 560 GCCTGGAGAGGCTGCATGAAATTCAGTTCTCTCCAGACAGCTTGGCGCATCTT 619
QY 541 TTGCCCCGTATGAGTTGATGACTTACGAGACGCAATTTGTAATTGGGACTTTGA 597
Db 620 TTGCCCCGTATGAGTTGATGACTTACGAGACGCAATTTGTAATTGGGACTTTGA 676

RESULT 2
US-09-966-880A-1
; Sequence 1, Application US/09966880A
; Patent No. US2002016473A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966, 880A
; PRIORITY FILING DATE: 2001-09-28
; PRIORITY APPLICATION NUMBER: PCT/JP00/01918
; PRIORITY FILING DATE: 2000-03-28
; PRIORITY APPLICATION NUMBER: JP 11-371382
; PRIORITY FILING DATE: 1999-12-27
; PRIORITY APPLICATION NUMBER: JP 11-178999
; PRIORITY FILING DATE: 1999-06-24
; PRIORITY APPLICATION NUMBER: JP 11-87192
; PRIORITY FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2440
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (93)...(686)
; FEATURE:
; NAME/KEY: 5'UTR
```

```
/ LOCATION: (1)...(92)
/ FEATURE:
/ NAME/KEY: 3'UTR
/ LOCATION: (690)...(2440)
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(2440)
/ OTHER INFORMATION: n = A,T,C or G
US-09-966-880A-1
```

Query Match 79.6%; Score 475.4; DB 9; Length 2440;
Best Local Similarity 87.3%; Pred. No. 4e-147;
Matches 521; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

```
QY 1 ATGACAGCCTCTGATGAAACCGAGAGATTCTTTACCAATTCAAAATATGCCCTGG 60
Db 93 ATGACAGCCTCTGATGAAACCGAGAGATTCTTTACCAATTCAAAATATGCCCTGG 152
QY 61 GCTAAGGATCGCGTGAAGACCTACCTGTGTAAGTAAGAGCGGTGACAGTGTACA 120
Db 153 GCTAAGGATCGCGTGAAGACCTACCTGTGTAAGTAAGAGCGGTGACAGTGTACA 212
QY 121 TCTTTTCTACTGACTTTGTGTATCTTCCGATTAAGAAAGCGTCCACCTGGAATTGCTC 180
Db 213 TCTTTTCTACTGACTTTGTGTATCTTCCGATTAAGAAAGCGTCCACCTGGAATTGCTC 272
QY 181 TTCTCCGCTACATCTCGGACTGGGACCTAGACCTGGCGCGCTGCTACCGCGTCACTGG 240
Db 273 TTCTCCGCTACATCTCGGACTGGGACCTAGACCTGGCGCGCTGCTACCGCGTCACTGG 332
QY 241 TTCACTCTGAGACCCCTGCTACGACTGTGCCGACATGTGGCCGACATTTCTGCGAGGG 300
Db 333 TTCACTCTGAGACCCCTGCTACGACTGTGCCGACATGTGGCCGACATTTCTGCGAGGG 392
QY 301 AACCCCAACTCACTGCTGAGGATCTTCAACCGCGCGCTTACTTCTGTGAGAGACGCAAG 360
Db 393 AACCCCAACTCACTGCTGAGGATCTTCAACCGCGCGCTTACTTCTGTGAGAGACGCAAG 452
QY 361 GCTGAGCCGAGGGGCTGCGCGGCTGCGACCGCGCGGGGTGCAAAATAGCATATGACC 420
Db 453 GCTGAGCCGAGGGGCTGCGCGGCTGCGACCGCGCGGGGTGCAAAATAGCATATGACC 512
QY 421 TTCAAGATATTTTATCTGCTGAGAACTTTGTGAGAAAACCATGAAGAACTTTCAAA 480
Db 513 TTCAAGATATTTTATCTGCTGAGAACTTTGTGAGAAAACCATGAAGAACTTTCAAA 572
QY 481 GCCTGGAGAGGCTGCATGAAATTCAGTTCTCTCCAGACAGCTTGGCGCATCTT 540
Db 573 GCCTGGAGAGGCTGCATGAAATTCAGTTCTCTCCAGACAGCTTGGCGCATCTT 632
QY 541 TTGCCCCGTATGAGTTGATGACTTACGAGACGCAATTTGTAATTGGGACTTTGA 597
Db 633 TTGCCCCGTATGAGTTGATGACTTACGAGACGCAATTTGTAATTGGGACTTTGA 689

RESULT 3
US-09-966-880A-10
; Sequence 10, Application US/09966880A
; Patent No. US2002016473A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966, 880A
; PRIORITY FILING DATE: 2001-09-28
; PRIORITY APPLICATION NUMBER: PCT/JP00/01918
; PRIORITY FILING DATE: 2000-03-28
; PRIORITY APPLICATION NUMBER: JP 11-371382
; PRIORITY FILING DATE: 1999-12-27
; PRIORITY APPLICATION NUMBER: JP 11-178999
; PRIORITY FILING DATE: 1999-06-24
; PRIORITY APPLICATION NUMBER: JP 11-87192
```

;; PRIOR FILING DATE: 1999-03-29
;; NUMBER OF SEQ ID NOS: 36
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 10
;; LENGTH: 6564
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-966-880A-10

Query Match 45.8%; Score 273.4; DB 9; Length 6564;
Best Local Similarity 99.6%; Pred. No. 1.3e-79;
Matches 274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 155 AGAAGCGCTGCGACGATGGAATGCTTCTCCGCTACATCTGGAGCTGGAGACCTAGACC 214
DB 2590 AGAAGCGCTGCGACGATGGAATGCTTCTCCGCTACATCTGGAGCTGGAGACCTAGACC 2649
QY 215 CTGACCGCTGCTACCGCGCTCACTGATTCACTCTGGAGCCCTGCTACGACTGTGCC 274
DB 2650 CTGACCGCTGCTACCGCGCTCACTGATTCACTCTGGAGCCCTGCTACGACTGTGCC 2709
QY 275 GACATGTGCGCGACTTTCTGCGAGGAAACCCCACTCACTGAGGATCTTACCGCGC 334
DB 2710 GACATGTGCGCGACTTTCTGCGAGGAAACCCCACTCACTGAGGATCTTACCGCGC 2769
QY 335 GCCTCTACTTCTGTGAGAGACCGCAAGGCTGAGCCCGGCGCTGCAACCGC 394
DB 2770 GCCTCTACTTCTGTGAGAGACCGCAAGGCTGAGCCCGGCGCTGCAACCGC 2829
QY 395 CCGGGGTGCAATATAGCATCATGACCTTCAAGAT 429
DB 2830 CCGGGGTGCAATATAGCATCATGACCTTCAAGAT 2864

RESULT 4

US-09-966-880A-35
; Sequence 35, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 11204
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-966-880A-35

Query Match 45.8%; Score 273.4; DB 9; Length 11204;
Best Local Similarity 99.6%; Pred. No. 1.8e-79;
Matches 274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 155 AGAAGCGCTGCGACGATGGAATGCTTCTCCGCTACATCTGGAGCTGGAGACCTAGACC 214
DB 7805 AGAAGCGCTGCGACGATGGAATGCTTCTCCGCTACATCTGGAGCTGGAGACCTAGACC 7864
QY 215 CTGACCGCTGCTACCGCGCTCACTGATTCACTCTGGAGCCCTGCTACGACTGTGCC 274
DB 7865 CTGACCGCTGCTACCGCGCTCACTGATTCACTCTGGAGCCCTGCTACGACTGTGCC 7924

QY 275 GACATGTGCGCGACTTTCTGCGAGGAAACCCCACTCACTGAGATCTTACCGCGC 334
DB 7925 GACATGTGCGCGACTTTCTGCGAGGAAACCCCACTCACTGAGATCTTACCGCGC 7984
QY 335 GCCTCTACTTCTGTGAGAGACCGCAAGGCTGAGCCCGGCGCTGCAACCGC 394
DB 7985 GCCTCTACTTCTGTGAGAGACCGCAAGGCTGAGCCCGGCGCTGCAACCGC 8044
QY 395 CCGGGGTGCAATATAGCATCATGACCTTCAAGAT 429
DB 8045 CCGGGGTGCAATATAGCATCATGACCTTCAAGAT 8079

RESULT 5

US-09-966-880A-13
; Sequence 13, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 271
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-966-880A-13

Query Match 45.4%; Score 271; DB 9; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.4e-79;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 AACGGCTGCAAGTGAATGCTTCTCCGCTACATCTGGAATGGAACCTAGACCT 216
DB 1 AACGGCTGCAAGTGAATGCTTCTCCGCTACATCTGGAATGGAACCTAGACCT 60
QY 217 GCGCGCTGCTACCGCGCTCACTGATTCACTCTGAGAGCCCTGCTACGACTGTGCC 276
DB 61 GCGCGCTGCTACCGCGCTCACTGATTCACTCTGAGAGCCCTGCTACGACTGTGCC 120
QY 277 CATGTGCGCGACTTTCTGCGAGGAAACCCCACTCACTGAGATCTTACCGCGC 336
DB 121 CATGTGCGCGACTTTCTGCGAGGAAACCCCACTCACTGAGATCTTACCGCGC 180
QY 337 CTCTACTTCTGTGAGAGACCGCAAGGCTGAGCCCGGCGCTGCAACCGC 396
DB 181 CTCTACTTCTGTGAGAGACCGCAAGGCTGAGCCCGGCGCTGCAACCGC 240
QY 397 GGGGTGCAATATAGCATCATGACCTTCAAG 427
DB 241 GGGGTGCAATATAGCATCATGACCTTCAAG 271

RESULT 6

US-09-966-880A-12
; Sequence 12, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE

FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966, 880A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-05-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 148
TYPE: DNA
ORGANISM: Homo sapiens
US-09-966-880A-12

Query Match 24.8%; Score 148; DB 9; Length 148;
Best Local Similarity 100.0%; Pred. No. 9, 6e-39;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 CCTCTGATGACCGGAGAGATTCTTACCAATTCAAAATGTCGGCTGGCTAAGG 68
DB 1 CCTTTATGACCGGAGAGATTCTTACCAATTCAAAATGTCGGCTGGCTAAGG 60
QY 69 TCGGCGTGAACCTACTGCTGCTAGTAGAAGGCGTGAAGCTGCTAATCTTTTC 128
DB 61 TCGGCGTGAACCTACTGCTGCTAGTAGAAGGCGTGAAGCTGCTAATCTTTTC 120
QY 129 ACTGACTTGTGTTATCTTGCAATAG 156
DB 121 ACTGACTTGTGTTATCTTGCAATAG 148

RESULT 7

US-10-460-923-1
Sequence 1, Application US/10460923
Publication No. US2004000955A1
GENERAL INFORMATION:
APPLICANT: MALIM, Michael H.
APPLICANT: SHEEHY, Ann M.
APPLICANT: HARRIS, Reuben S.
APPLICANT: BISHOP, Kate N.
APPLICANT: NEUBERGER, Michael S.
APPLICANT: GADDIS, Nathan C.
APPLICANT: SIMON, James H.M.
TITLE OF INVENTION: DNA Deamination Mediates Innate Immunity to Retroviral Infection
FILE REFERENCE: 2253-74380
CURRENT APPLICATION NUMBER: US/10/460, 923
CURRENT FILING DATE: 2003-06-13
PRIOR APPLICATION NUMBER: US 60/388, 513
PRIOR FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: US 60/472, 952
PRIOR FILING DATE: 2003-05-23
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1155
TYPE: DNA
ORGANISM: Homo sapiens
US-10-460-923-1

Query Match 19.8%; Score 118.4; DB 15; Length 1155;
Best Local Similarity 58.9%; Pred. No. 2, 2e-28;
Matches 225; Conservative 0; Mismatches 151; Indels 6; Gaps 1;

QY 158 ACGGTCGACAGTGAATTCCTCCGCTACATCTCGAGCTGGAGCTAGACCTG 217
DB 761 AAGGCGCATGCAAGCTGCTCTTCTGAGCGATTCCTTTGGAAGCTGACCTG 820
QY 218 GCGGCTGCTACCGGCTACCTGCTTCACTCTCTGAGCCCTGCTAGACTGCGCCGAC 277

DB 821 ACCGAGACTACAGGTTACCTGCTCACTCTGAGACCCCTCTTACGCTGCGCCAG 880
QY 278 ATGAGCGGACTTCTGCGAGGAGACCCCACTCAGTCTGAGATCTTACCGCGCC 337
DB 861 AATGCGTAATTCATTTCAAAACCAACGCTGAGCCCTGCTGATCTTACGCGCCGA 940
QY 338 TCTACTTCTGAGAGACCGAAGGCTAGGCCGAGGGCTGCGCGGCTGACCGCGCCG 397
DB 941 TCTA-----TGATGATCAAGAAAGATGTCAGAGGGGCTGCGCACCTGCGAGGCTG 994
QY 398 GGGTGCAATGACCATGATGACTTCAAGATTATTTTACTGTGAAATCTTTTAG 457
DB 995 GGGCCAAATTCATATGATGACATACGATGATTTAGACTGTGAGACCTTTGTGG 1054
QY 458 AAACCATGAGAAACCTTCAAGCCGAGGAGGCTGCTGAAATTCAGTGTCTCT 517
DB 1055 ACCACGAGGATGTCCTTCCAGCCCTGGATGAGACTAGATGACACAGCAAGACTGA 1114
QY 518 CCAAGAGCTTGGCGCATCT 539
DB 1115 GTGGAGGCTGCGGCAATCT 1136

RESULT 8

US-09-729-674-173
Sequence 173, Application US/09729674
Patent No. US2001003935A1
GENERAL INFORMATION:
APPLICANT: JACOBS, Kenneth
APPLICANT: MCCOY, John M.
APPLICANT: LAVAILLE, Edward R.
APPLICANT: COLLINS-RACIE, Lisa A.
APPLICANT: EVANS, Cheryl
APPLICANT: MERBERG, David
APPLICANT: TREACY, Maurice
APPLICANT: AGOSTINO, Michael J.
APPLICANT: STEININGER II, Robert J.
APPLICANT: SPALDING, Vikki
APPLICANT: WONG, Gordon G.
APPLICANT: CLARK, Hilary
APPLICANT: FECHTEL, Kim
TITLE OF INVENTION: GENETICS INSTITUTE, INC.
FILE REFERENCE: 6055-64X
CURRENT APPLICATION NUMBER: US/09/729, 674
CURRENT FILING DATE: 2000-12-04
PRIOR APPLICATION NUMBER: 09/539, 330
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 283
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 173
LENGTH: 1534
TYPE: DNA
ORGANISM: Homo sapiens
US-09-729-674-173

Query Match 19.8%; Score 118.4; DB 9; Length 1534;
Best Local Similarity 58.9%; Pred. No. 2, 5e-28;
Matches 225; Conservative 0; Mismatches 151; Indels 6; Gaps 1;

QY 158 ACGGTCGACAGTGAATTCCTCCGCTACATCTCGAGCTGGAGCTAGACCTG 217
DB 811 AAGGCGCATGCAAGCTGCTCTTCTGAGCGATTCCTTTGGAAGCTGACCTG 870
QY 218 GCGGCTGCTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 277
DB 871 ACCAGACTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 930
QY 278 ATGAGCGGACTTCTGCGAGGAGACCCCACTCAGTCTGAGATCTTACCGCGCC 337
DB 931 AATGCGTAATTCATTTCAAAACCAACGCTGAGCTGCTGCTGCTGCTGCTGCTG 990

| QY | 338 | CTCACTTCTGTGAGGACCGCAAGGCTGAGCCCGAAGGGGTGTGGGGCTGCACCGCGCG | 397 |
|----|------|--|------|
| QY | 991 | TCCTA-----TCATATCAAGGAABATGTCAAGAGGGGTGTGGCACCTGTGGCTGAGGCTG | 1044 |
| QY | 398 | GGGTGCAATTAAGCCATCATGACCTTCAAGATTTATTTTACTGCTGAATACTTTTGTAG | 457 |
| Db | 1045 | GGGCCAAATTTTCATTAATGACATCACTAGTGATTTAAACACTGTGGGACACCTTTGTGG | 1104 |
| QY | 458 | AAATCCATGAAAGAACTTTCAAGGCTGTGGAAAGGCTGCAATGAAATTCAGTTGCTCTCT | 517 |
| Db | 1105 | ACCAACCAAGAGATGTCCTCTTCCAGCCCTGTGGATGACTAGATGAGCAAGCCAAAGACTGA | 1164 |
| QY | 518 | CCAGACAGCTTGGCGGCATCCT | 539 |
| Db | 1165 | GTGGGAGGCTGTGGGCCATTTCT | 1186 |

```

US-09-925-300-699
RESULT 9
US-09-925-300-699
Sequence 699, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
APPLICANT: Steve Ruben,
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 699
LENGTH: 987
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-300-699

```

| | | | | |
|-----------------------|--------------|--------------------|-----------------|-------------|
| Query Match | 19.2%; | Score 114.8; | DB 9; | Length 987; |
| Best Local Similarity | 53.9%; | Pred. No. 3.1e-27; | | |
| Matches 289; | Conservative | 0; | Mismatches 232; | Indels 15; |
| | | | | Gaps 2; |

QY 16 ATGAACCGAGGAAAGTTTCTTTACCAATTCAAAAATGTCCCGTGGCGTCAACAGGTCCGGCGT 75

Db 132 ATGATACAGGCAACATTCTTACTTCACATTAAAACTATGGGAAGCCACAGATCCGAGAC 191

QY 76 GAGACCTACCTGTGCTACAGNATGGAAGAGCGCTGACAGNGCTACATCCTTTCTACTGGAC 135

Db 192 GAAACTTGGCTGTGCTTCAACCGTGAAGGTATAAAGCCCGCTCAGTTGTCTCCTGGAG 251

QY 136 TTTGGTATCTTGGCAA-----TAAGAAGCGCTGCACGNTGAAATTGCTTTC 183

Db 252 ACGGGCGTCTTCCGAAACCAAGTGGATTCTGAGCCCATGTGCTATGACAGAAAGTGCTTC 311

QY 184 CTCGCGTACATCTCGGACTGGAGCTGAGACCTCGCGCTGTAACCGCGTCACTCGTGT 243

Db 312 CTCCTCTGGTCTCGACACACATACCTGTCTCTTAACAAAGTACCAAGTACACTGGTAC 371

QY 244 ACCTCTGGAAGCCCTGCTACGACTGTGCCCCGACATGTGTGCGCACTTCTTCCGAGGAC 303

Db 372 ACATCTTGGAGCCCTTGGCCGACACTGTGAGGGGAGGTGCGCACTTCTTGGCCAGGAC 431

QY 304 CCCAACCCTCAGTCTGAGATCTTCAACCGCGCGCTCTACTTCTGTGAGAACCGCGAAGCT 353

Db 432 AGCAAGCTGAATCTCACACATCTTCAACCGCGCGCTCTACTTCTCAAGTATC---CATGT 488

QY 364 GAGCCCGAGGGGCTCGCGCGGCTGCACCGCGCCGGGCTGCAATATGCAATCATGACCTTC 423

Db 489 TACCAAGAGGGGCTCCGACAGCTGAATCAGAAAGGGGTGCTGTGAGATCATGACTAT 548

QY 424 AAGAGTATTTTTACTGCTGGAATACTTTTGTAGAAAACATGAAAGACTTTTAAAGCC 483

Db 549 GAAAGATTTTAAATATTTGCGAAAACTTTGTGTCAATGATTAATGAGCCATTAAAGCTT 608

Oy 484 TGGAAAGGCGTCGATGAAATTCAGTTCGTCTCTCCAGACAGCTTCGGCGCATCTT 539

Db 609 TGGAAAGGATTAATAAACCAACTTTGCACCTTCGAAAAAGAGCCTACGGAGAGCTT 664

RESULT 10
US-10-098-841-268
; Sequence 268, Application US/10098841
; Publication No. US20020197679A1
CURRENT INFORMATION.

```

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Yindod
APPLICANT: Xu, Chongjun
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Duntui
APPLICANT: Wang, Zhiwei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Qian, Xiaohong B.
APPLICANT: Dimaenc, Radje T.
TITLE OF INVENTION: No. US20020197679a1el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2
CURRENT APPLICATION NUMBER: US/10/098,841
CURRENT FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 331
SOFTWARE: pf_FL_genes Version 1.0
SEQ ID NO 268
LENGTH: 1143
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (128)..(700)
US-10-098-841-268

```

| | | | | | | | |
|-----------------------|-------|--------------|---------|------------|-----|--------|------|
| Query Match | 19.2% | Score | 114.8 | DB | 13 | Length | 1143 |
| Best Local Similarity | 53.9% | Pred. No. | 3.4e-27 | | | | |
| Matches | 289 | Conservative | 0 | Mismatches | 234 | Indels | 15 |
| | | | | | | Gaps | 2 |

QY 16 ATGAACCGAGAGAGAGTTTCTTACCAATTCAAAATATGCGCTGGCGTAAAGGGTCGGCGT 75

Db 161 ATGTATCCAGGCACATTTCTACTTCCATTTAAATTAACCTATGGAAAGCCACGATCCGAAAC 22

QY 76 GAGACCTACCTGTCTGTACTGTAAGTGAAGAGCGGTGACAGTCTACATCTTTTCACTGAC 13

Db 221 GAACACTGGCTGTGCTTCAACGCTGGAAGGTATTAAGACCCGCTCAGTTGTCTCTGGAAG 28

QY 136 TTGGTATATCTTGCAAA-----TAAAGACGGCTGCCAGTGGAAATGCTCTTC 183

Db 281 ACGGGCGTCTTCGAAACACAGGTGATTTCTGAGCCCATTTGATCGAGAAAGGTGCTTC 34

QY 184 CTCGCGTAAATCTCCGGAACGTGGACCTTAACCCGTGGCGCTCTACCGCGTCACTGGTTC 243

Db 341 CTCCTTGGTTCTGCGACACGACATCTGTCTCTTAACAAGATACAGGTGACCTGTGATC 40

QY 244 AACTCTCTGAGACCCCTCTGCTACGACTGTGCGCGACATGTGGCCGACTTTTGTCCGAGGGAAC 303

DB 401 ACATTTGAGACCCCTTCCGACGCTTGTGAGGAGGCTGCGGAGTTCTTCCGACGAC 460
QY 304 CCCAACCTCAGTCTGAGGATCTTTACCGCGGCTCTTACTTTGTGAGGACCGGAGGCT 363
DB 461 AGCAACGTAATCTACCATCTTCCACGCGGCTCTTACTTCACTATC---CATGT 517
QY 364 GAGCCCGAGGCGGCTCGCGGCTGACCGCGGCTGCAAAATGACCATCATGACCTTC 423
DB 518 TACCAAGAGGCGGCTCGCGGCTGAGTCAAGAGGCGGCTGCTGTGAGATCATGACCTAT 577
QY 424 AAAAGTATTTTACTGCTGGAATCTTTGTGAAACCATGAAAGAACTTGAAGCC 483
DB 578 GAAGATTTTAAATATTTGTTGGGAAAATTGTGTAATGATGAGCCATTCAGCT 637
QY 484 TGGGAGGCGCTCATGAAATTCAGTCTCTCTCCAGACAGCTTGGCGATCCT 539
DB 638 TGGAAAGGATTAACCACTTTCGACTTCTGAAAAGAGCTACGGGAGAGTCT 693

RESULT 11

US-09-966-880a-14
Sequence 14, Application US/09966880A
Patent No. US2002016473A1
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966, 880A
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 116
TYPE: DNA
ORGANISM: Homo sapiens
US-09-966-880a-14

Query Match 19.2%; Score 114.4; DB 9; Length 116;
Best Local Similarity 99.1%; Pred. No. 1.3e-27;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 428 ATTATTTTACTGCTGGAATCTTTGTGAAACCATGAAAGAACTTGAAGCTCGG 487
DB 1 ATTATTTTACTGCTGGAATCTTTGTGAAACCATGAAAGAACTTGAAGCTCGG 60
QY 488 AAGGCGTGAATGAATTCAGTCTCTCCAGACAGCTTGGCGCATCTTTTG 543
DB 61 AAGGCGTGAATGAATTCAGTCTCTCCAGACAGCTTGGCGCATCTTTTG 116

RESULT 12

US-09-880-107-3282
Sequence 3282, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379

PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3282
LENGTH: 1348
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 U03891
NAME/KEY: unsure
LOCATION: (1) (1348)
OTHER INFORMATION: n = a or c or g or t
US-09-880-107-3282

Query Match 18.5%; Score 110.4; DB 9; Length 1348;
Best Local Similarity 59.3%; Pred. No. 1.1e-25;
Matches 230; Conservative 0; Mismatches 146; Indels 12; Gaps 2;

QY 158 ACCGCTGCCACGATGATTTCTTCTCCGCTACATCTGAGCTGAGACTAGACCTTG 217
DB 275 ACGGCGGCAATGGGAGCTGCGCTTCTTGACCTGGTCTTTGACATTTGACCCGG 334
QY 218 GCGCTGCTACCGCGCTACCTGTTTCACTCTCTGAGCCCGCTAC-----GACTGTG 271
DB 335 CCCAGATCTACAGAGTCACTTGTTCATCTCTGAGCCCTGCTTCTGAGGCTG 394
QY 272 CCGGACATGAGGCGGCTTCTGAGGAGAACCCCACTCACTGAGATTTTACCG 331
DB 395 CCGGAGATGCTGCTGCTTCTTCAAGAAACACACTGAGACTGCGCATCTTGCTG 454
QY 332 CCGGCTCTACTTCTGAGGACCGGAGGCTGAGGCGGCTGCGGCTGACCG 391
DB 455 CCGGCACTA-----TGATTAAGACCCCTATATAAGAGGCGCTGCAATGCTGCGG 508
QY 392 GCGCGGCGGCTGCAAAATGACCATATGACTTCAAGATTTTATCTGCTGGAATCT 451
DB 509 ATGCTGGGCGGCAAGCTCTCATCATGACCTACGATTAATTTAAGCATCTGAGACCT 568
QY 452 TTGTAAGAAACCATGAAAGAACTTCAAGGCTGAGGAGGCTGCAAAATTCAGTTC 511
DB 569 TTGTGACCAACGAGGATGTCCTTCAAGGCTTGGAGTGAATGAGCAGACGCAAG 628
QY 512 GTCTCTCCAGACAGCTTGGCGCATCCT 539
DB 629 CCTGAGTGGAGGCTGCGGCGCATTC 656

RESULT 13

US-10-247-671-62
Sequence 62, Application US/10247671
Publication No. US20030194721A1
GENERAL INFORMATION:
APPLICANT: Mikita, Thomas
APPLICANT: Shiffman, Dov
APPLICANT: Porter, Gordon, J.
APPLICANT: Kaser, Matthew R.
TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
FILE REFERENCE: PA-0050 US
CURRENT APPLICATION NUMBER: US/10/247,671
CURRENT FILING DATE: 2002-09-18
PRIOR APPLICATION NUMBER: 60/323,784
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 186
SOFTWARE: PERL Program
SEQ ID NO 62
LENGTH: 823
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: incyte ID No. US20030194721A1 135626CBI

US-10-247-671-62

Query Match 18.0%; Score 107.2; DB 14; Length 823;
Best Local Similarity 58.8%; Pred. No. 9.6e-25;
Matches 228; Conservative 0; Mismatches 148; Indels 12; Gaps 2;

QY 158 ACGGCTGCCAGCTGGAATTGCTTCTCCGCTACATCTCGAGCTGGACCTAGACCTG 217
DB 115 ACGGCTGCCAGCTGGAATTGCTTCTCCGCTACATCTCGAGCTGGACCTAGACCTG 174
QY 218 GCGGCTGCTACCGCGCTGACCTGCTTCCCTCGAGCCCTGCTAC-----GACTGTG 271
DB 175 CCGAGATCTACAGGGCTCACTTGGTTATCTCCTGGAGCCCTGCTTCTCGGGCTGTG 234
QY 272 CCGGACATGTGCGCGCACTTCTGCGAGGAAACCCCACTCACTGCTGAGATCTTACCG 331
DB 235 CCGGGGAACTGGCGGTGCTTCTTCAAGAGAACACACACTGAGCTGGCATCTTGGCTG 294
QY 332 CCGGCTCTACTTCTGTGAGAACCGCAAGCTGAGCGCGAGGGGCTGGCGGCTGCAAC 391
DB 295 CCGGATCTATGATTAATGACCCCTCTATATAGAGGCGCTGCAAAATGCTGCGGG 348
QY 392 GCGCGCGGGTGCMAATAGCATCATGACCTTCAAGATTATTTTACTGCTGGAATCT 451
DB 349 ATGCTGGGGCCCAAGTCTCATCATGACCTTCAAGATGAGTTGAGTACTGCTGGACACT 408
QY 452 TTGTAGAAAACCATGAAAGAACTTTCMAAGCTTGGAGGGCTGCAATGAAATTCAGTTC 511
DB 409 TTGTGACCGCCAGGAGATGCTCCCTTCACGCTGGGATGAGTAAAGAGACACAGCAAG 468
QY 512 GTCTCTCCAGACAGCTTGGCGGATCCT 539
DB 469 CCTGAGTGGAGGCTGCGGCGCATCT 496

RESULT 14

US-10-131-827-8890
; Sequence 8890, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
; APPLICANT: Mohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
; TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES
; FILE REFERENCE: 506612000120
; CURRENT APPLICATION NUMBER: US/10/131,827
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US 10/006,290
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/296,764
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 9090
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8890
; LENGTH: 1008
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-827-8890

Query Match 18.0%; Score 107.2; DB 15; Length 1008;
Best Local Similarity 58.8%; Pred. No. 1.1e-24;
Matches 228; Conservative 0; Mismatches 148; Indels 12; Gaps 2;

QY 158 ACGGCTGCCAGCTGGAATTGCTTCTCCGCTACATCTCGAGCTGGACCTAGACCTG 217
DB 252 ACGGCTGCCAGCTGGAATTGCTTCTCCGCTACATCTCGAGCTGGACCTAGACCTG 311
QY 218 GCGGCTGCTACCGCGCTGACCTGCTTCCCTCGAGCCCTGCTAC-----GACTGTG 271
DB 312 CCGAGATCTACAGGGCTCACTTGGTTATCTCCTGGAGCCCTGCTTCTCGGGCTGTG 371

QY 272 CCGGACATGTGCGCGACTTCTGCGAGGAGACCCCACTCACTGAGTGAAGATCTTCAACG 331
DB 372 CCGGGAAGTGCCTGCTTCTTCAAGAGAAACACACAGTGAAGCTGCCGATCTTGTGCTG 431
QY 332 CCGGCTCTACTTCTGTGAGAGACCGCAAGGCTGAGGCCCGAGGGCTGCGCGGCTGACCC 391
DB 432 CCGGATCTATGATTAATGACCCCTCTATATAGAGGCGGCTGCAAAATGCTCGGG 485
QY 392 GCGCGGGGGTGCMAATAGCATCATGACCTTCAAGATTATTTTACTGCTGGAATCTT 451
DB 486 ATGCTGGGGCCCAAGTCTCATCATGACCTTCAAGATGAGTTGAGTCTGCTGGACACT 545
QY 452 TTGTAGAAAACCATGAAAGAACTTTCMAAGCCCTGGAGGGCTGCAATGAAATTCAGTTC 511
DB 546 TTGTGACCGCCAGGAGTGTCTTCTTCAAGCTTGGAGTGAAGTGAAGAGACACAGCAAG 605
QY 512 GTCTCTCCAGACAGCTTGGCGGATCCT 539
DB 606 CCTGAGTGGAGGCTGCGGCGCATCT 633

RESULT 15

US-09-822-830A-359
; Sequence 359, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fachel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402
; CURRENT APPLICATION NUMBER: US/09/822,830A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,604
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 359
; LENGTH: 2151
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-830A-359

Query Match 17.5%; Score 104.4; DB 9; Length 2151;
Best Local Similarity 58.4%; Pred. No. 1.4e-23;
Matches 223; Conservative 0; Mismatches 151; Indels 8; Gaps 2;

QY 158 ACGGCTGCCAGCTGGAATTGCTTCTCCGCTACATCTCGAGCTGGACCTAGACCTG 217
DB 1450 AAGGCGCCCATGACAGACTGTCTTCTCGAGCTGATTCCTTTTGAAGCTGAGACTG 1509
QY 218 GCGGCTGCTACCGCGTACCTGTTCACTCTCGAGCCCGCTACGACTGAGCCGAC 277
DB 1510 ACGAGACTACAGGGTATCTGCTTCACTCTCGAGCCCGCTTCACTGAGTGGCCAGG 1569
QY 278 ATGAGCGCACTTCTGCGAGGAGACCCCACTCACTGAGTGAAGATCTTCAACCGCGGC 337
DB 1570 AATGCTTAATTCATTTCAAAAACAAACAGTGAAGCTGAGCTTCACTGAGCCGCA 1629
QY 1630 TCTA-----TGATGATCAAGAAAGTCAAGAGGGGCTGCGACCTGCGAGGCTG 1683
QY 398 GGGTGCAATAGCATATGACCTTCAAGATATTTTACTGCTGGAATATCTTTGTAG 457
DB 1684 GGGCAAAATTTCAATATGACATACAGTGAATTTAAGCACTGCTGAGACACTTTGG--G 1741

Mon Mar 15 09:28:01 2004

us-09-966-880a-7_copy_80_676.rnpb

Page 8

| QY | 458 | AAAACCATGAAAGAAACCTTCAAGCGTGGGAAGGGCTGCATGAAATTCAGTTGCTCT | 517 |
|----|------|---|------|
| DB | 1742 | ACACACCAAGGATGTCTCCACGCCCTGGATGACTAGATGAGACAAGCCAAAGCTGA | 1801 |
| QY | 518 | CCAGACAGCTTCGGCGCATCCT | 539 |
| DB | 1802 | GTGGAGGCTCGCGCCATTC | 1823 |

```
Search completed: March 12, 2004, 12:26:29
Job time : 250 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2004, 17:34:19 ; Search time 22 Seconds
(without alignments)
464.634 Million cell updates/sec

Title: US-09-966-880A-8
Perfect score: 1086
Sequence: 1 MDSLMNRKFLYFKNVRW.....ILFLYVDDLRDAPRTLGI 198

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum March 0%
Maximum March 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--|
| 1 | 369.5 | 34.0 | 190 | 1 | US-08-816-241-1 Sequence 1, Appl |
| 2 | 369.5 | 34.0 | 190 | 3 | US-09-128-395-1 Sequence 1, Appl |
| 3 | 271 | 25.0 | 222 | 1 | US-08-687-895-1 Sequence 1, Appl |
| 4 | 271 | 25.0 | 222 | 2 | US-09-040-482-1 Sequence 1, Appl |
| 5 | 245 | 22.6 | 116 | 1 | US-08-687-895-3 Sequence 3, Appl |
| 6 | 245 | 22.6 | 116 | 1 | US-08-816-241-3 Sequence 3, Appl |
| 7 | 245 | 22.6 | 116 | 1 | US-09-040-482-3 Sequence 3, Appl |
| 8 | 245 | 22.6 | 116 | 2 | US-09-128-395-3 Sequence 3, Appl |
| 9 | 225 | 20.7 | 275 | 4 | US-09-079-030-224 Sequence 2, Appl |
| 10 | 216.5 | 19.9 | 236 | 1 | US-08-687-895-4 Sequence 4, Appl |
| 11 | 216.5 | 19.9 | 236 | 1 | US-08-816-241-4 Sequence 4, Appl |
| 12 | 216.5 | 19.9 | 236 | 2 | US-09-040-482-4 Sequence 4, Appl |
| 13 | 216.5 | 19.9 | 236 | 2 | US-09-128-395-4 Sequence 4, Appl |
| 14 | 210 | 19.3 | 236 | 1 | US-08-158-682A-4 Sequence 4, Appl |
| 15 | 198 | 18.2 | 229 | 1 | US-08-158-682A-2 Sequence 2, Appl |
| 16 | 198 | 18.2 | 229 | 1 | US-08-015-203-2 Sequence 2, Appl |
| 17 | 198 | 18.2 | 229 | 1 | US-08-687-895-5 Sequence 5, Appl |
| 18 | 198 | 18.2 | 229 | 1 | US-08-816-241-5 Sequence 5, Appl |
| 19 | 198 | 18.2 | 229 | 2 | US-09-040-482-5 Sequence 5, Appl |
| 20 | 198 | 18.2 | 229 | 2 | US-09-128-395-5 Sequence 5, Appl |
| 21 | 81 | 7.5 | 440 | 4 | US-09-092-315-3 Sequence 3, Appl |
| 22 | 81 | 7.5 | 440 | 4 | US-09-733-524A-3 Sequence 3, Appl |
| 23 | 80.5 | 7.4 | 451 | 4 | US-09-107-532A-6444 Sequence 6444, Ap |
| 24 | 80 | 7.4 | 476 | 4 | US-09-092-315-5 Sequence 5, Appl |
| 25 | 79.5 | 7.3 | 476 | 4 | US-09-733-524A-5 Sequence 5, Appl |
| 26 | 79.5 | 7.3 | 271 | 4 | US-09-252-991A-24319 Sequence 24319, A |
| 27 | 77.5 | 7.1 | 402 | 6 | 5352575-5 Patent No. 5352575 |

| | | | | | |
|----|------|-----|------|---|--|
| 28 | 75.5 | 7.0 | 343 | 1 | US-08-180-209B-56 Sequence 56, Appl |
| 29 | 75.5 | 7.0 | 343 | 3 | US-08-474-853-56 Sequence 56, Appl |
| 30 | 75.5 | 7.0 | 343 | 3 | US-09-166-205B-56 Sequence 56, Appl |
| 31 | 75.5 | 7.0 | 343 | 5 | PCT-US94-02629-56 Sequence 5, Appl |
| 32 | 75.5 | 7.0 | 347 | 4 | US-09-806-658-10 Sequence 10, Appl |
| 33 | 73.5 | 6.8 | 533 | 1 | US-08-488-305A-6 Sequence 6, Appl |
| 34 | 72.5 | 6.7 | 366 | 6 | 5470718-4 Patent No. 5470718 |
| 35 | 72 | 6.6 | 487 | 4 | US-09-489-039A-9810 Sequence 9810, Ap |
| 36 | 72 | 6.6 | 1094 | 2 | US-08-680-326-40 Sequence 40, Appl |
| 37 | 72 | 6.6 | 1400 | 1 | US-08-080-255-7 Sequence 7, Appl |
| 38 | 72 | 6.6 | 1400 | 3 | US-08-465-713-7 Sequence 7, Appl |
| 39 | 72 | 6.6 | 1400 | 5 | PCT-US93-05857-7 Sequence 7, Appl |
| 40 | 72 | 6.6 | 3969 | 3 | US-08-061-376-5 Sequence 5, Appl |
| 41 | 71 | 6.5 | 350 | 4 | US-09-489-039A-10384 Sequence 10384, A |
| 42 | 71 | 6.5 | 351 | 4 | US-09-252-991A-23740 Sequence 23740, A |
| 43 | 71 | 6.5 | 860 | 4 | US-09-936-989A-2 Sequence 2, Appl |
| 44 | 71 | 6.5 | 925 | 4 | US-09-936-989A-6 Patent No. 5475095 |
| 45 | 70.5 | 6.5 | 528 | 6 | 5475095-2 Patent No. 5475095 |

ALIGNMENTS

RESULT 1
US-08-816-241-1
Sequence 1, Application US/08816241
Patent No. 5804185
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL RNA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: Inocyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASTESQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/816,241
FILING DATE: Filed Herewith
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 190 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSTTT09
CLONE: 1646823
US-08-816-241-1
Query Match 34.0%; Score 369.5; DB 1; Length 190;
Best Local Similarity 44.9%; Pred. No. 2.5e-35;
Matches 79; Conservative 24; Mismatches 64; Indels 9; Gaps 4;


```

NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,482
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/687,895
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0109 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 222 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: MUSCNOT1
CLONE: 57953
US-09-040-482-1

Query Match 25.0%; Score 271; DB 2; Length 222;
Best Local Similarity 35.8%; Pred. No. 1.1e-23;
Matches 63; Conservative 25; Mismatches 76; Indels 12; Gaps 5;

Cy 11 FLVYFKVNRKGRERETLYLVYKRDATSFLDFGLYLNKRGCHVELLFLRYSDMD 69
Db 52 FKFGFRRVYSSGNKTFLLCYVVEAGQGGQVQASRGYLEDHAAAHAEAFNTILP-A 110
Cy 70 LDPRCRRVMTFTNSCYOCARIVADFLGPNPLSLRTALYPCEDKAEPE--GL 126
Db 111 FDPALRYNVMTWYSSSCAACADRIKXTLSKTKRLRLILVGRLFMWE---EPEIQAL 166
Cy 127 RLHRAGVOIAIMTFKDYFCOMTFVENHE---RTFKAMEGLHNSVRLSRQLRRI 179
Db 167 KKLREAGCKLRIVKQDFEYVWQNFVEGEGESKAFQPMEDIOENFLYREKLADI 222

RESULT 5
US-08-687-895-3
Sequence 3, Application US/08687895
Patent No. 5747319
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Hawkins, Phillip R.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: A NOVEL HUMAN MRNA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,895
FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0109 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 436941
US-08-687-895-3

Query Match 22.6%; Score 245; DB 1; Length 116;
Best Local Similarity 46.4%; Pred. No. 5.1e-21;
Matches 51; Conservative 17; Mismatches 32; Indels 10; Gaps 3;

Cy 76 YRVMTFTSWSPCYD--CARHVADELFGPNLSLIFPARLYFCEDRAEP---EGLRLH 130
Db 7 YRVMTFTSWSPCYD--CARHVADELFGPNLSLIFPARLYFCEDRAEP---EGLRLH 130
Cy 131 RAGVOIAIMTFKDYFCOMTFVENHRTFKAMEGLHNSVRLSRQLRRI 180
Db 62 DAGAQSINTYDFEFCYMDTFYRQGCFFQPMDELHESQALSGRLAIL 111

RESULT 6
US-08-816-241-3
Sequence 3, Application US/08816241
Patent No. 5804185
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL RNA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/816,241
FILING DATE: Filed Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0239 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555

```

TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 436941
US-09-816-241-3

Query Match 22.6%; Score 245; DB 1; Length 116;
Best Local Similarity 46.4%; Pred. No. 5.1e-21;
Matches 51; Conservative 17; Mismatches 32; Indels 10; Gaps 3;

QY 76 YRVWTFWSPCYD--CARHVDPLRGPNLSLIFTRALYFCEDRKAEP---EGLRLH 130
DB 7 YRVWTFWSPCYD--CARHVDPLRGPNLSLIFTRALYFCEDRKAEP---EGLRLH 130
QY 131 RAGVOIATMTFQDYFCNNTFVENHRTFKAMEGLHNSVLSRQLRIL 180
DB 62 DAGAQSIVTIDFECYCMDFVYRQGCPCFQMDGLEHNSQALSGRLRAIL 111

RESULT 7
US-09-040-482-3
Sequence 3, Application US/09040482
Patent No. 5916556
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Hawkins, Phillip R.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: A NOVEL HUMAN MRNA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,482
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/667,895
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0109 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 436941
US-09-040-482-3

Query Match 22.6%; Score 245; DB 2; Length 116;

Best Local Similarity 46.4%; Pred. No. 5.1e-21;
Matches 51; Conservative 17; Mismatches 32; Indels 10; Gaps 3;

QY 76 YRVWTFWSPCYD--CARHVDPLRGPNLSLIFTRALYFCEDRKAEP---EGLRLH 130
DB 7 YRVWTFWSPCYD--CARHVDPLRGPNLSLIFTRALYFCEDRKAEP---EGLRLH 130
QY 131 RAGVOIATMTFQDYFCNNTFVENHRTFKAMEGLHNSVLSRQLRIL 180
DB 62 DAGAQSIVTIDFECYCMDFVYRQGCPCFQMDGLEHNSQALSGRLRAIL 111

RESULT 8
US-09-128-395-3
Sequence 3, Application US/09128395
Patent No. 6087108
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL RNA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/128,395
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/816,241
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0239 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 436941
US-09-128-395-3

Query Match 22.6%; Score 245; DB 3; Length 116;
Best Local Similarity 46.4%; Pred. No. 5.1e-21;
Matches 51; Conservative 17; Mismatches 32; Indels 10; Gaps 3;

QY 76 YRVWTFWSPCYD--CARHVDPLRGPNLSLIFTRALYFCEDRKAEP---EGLRLH 130
DB 7 YRVWTFWSPCYD--CARHVDPLRGPNLSLIFTRALYFCEDRKAEP---EGLRLH 130
QY 131 RAGVOIATMTFQDYFCNNTFVENHRTFKAMEGLHNSVLSRQLRIL 180
DB 62 DAGAQSIVTIDFECYCMDFVYRQGCPCFQMDGLEHNSQALSGRLRAIL 111

RESULT 9
US-09-079-030-224

; Sequence 224, Application US/09079030
; Patent No. 6635623
; GENERAL INFORMATION:
; APPLICANT: Guevera, Jr., Juan G.
; APPLICANT: Hoogeveen, Ron C.
; APPLICANT: Moore, Paul J.
; TITLE OF INVENTION: LIPOPROTEINS AS NUCLEIC ACID DELIVERY
; TITLE OF INVENTION: VECTORS FOR TRANSECTION OF EUKARYOTIC CELLS
; NUMBER OF SEQUENCES: 229
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,030
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McMillian, Nabeela R.
; REGISTRATION NUMBER: P-43,363
; REFERENCE/DOCKET NUMBER: ARAG:003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 224:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 275 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-09-079-030-224

Query Match 20.7%; Score 225; DB 4; Length 275;
Best Local Similarity 37.4%; Pred. No. 3.7e-18;
Matches 49; Conservative 26; Mismatches 48; Indels 8; Gaps 4;

Db 24 RRENYLCYVVKRRDSSATSLDFGYLANKGC-HVELLFL-RYISDWDIDPGRCYRVTFW 81
72 RKACCLVEIKWGNSSKTRRS-----GKNTTNHVEYNFLEKLTSGRLGSPSCSIITWF 126

Qy 82 TSMSPCYDCARHVAADFLRGNPNLSLRIFTARLYFCEDRKAPGRLRLHRAVQIAIMTF 141
127 LSNMPCWECMAIRFELSHQHPGVTLLIFVARLFGHMDR-NRQGLDVLTVSGTVAVMEV 185

Qy 142 KDIFYCWNFTV 152
186 SEVCYCWENFTV 196

Db 186 SEVCYCWENFTV 196

RESULT 10
US-08-687-895-4
; Sequence 4, Application US/08687895
; Patent No. 5747319
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: A NOVEL HUMAN MRNA EDITING ENZYME
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.

; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,895
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0109 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 1177798
; US-08-687-895-4

Query Match 19.9%; Score 216.5; DB 1; Length 236;
Best Local Similarity 34.9%; Pred. No. 3e-17;
Matches 44; Conservative 31; Mismatches 42; Indels 9; Gaps 4;

Qy 35 RDSATSSSLDFGYLR-----NRGC-HVELLFL-RYISDWDIDPGRCYRVTFWTSWSP 86
33 RKACCLVEIKWGNSSKTRRS-----GKNTTNHVEYNFLEKLTSGRLGSPSCSIITWF 126

Db 87 CYDCARHVAADFLRGNPNLSLRIFTARLYFCEDRKAPGRLRLHRAVQIAIMTFKRYFY 146
93 CWECQAIRFELSHQHPGVTLLIFVARLFGHMDR-NRQGLDVLTVSGTVAVMEV 185

Qy 147 CWNFTV 152
152 CWNFTV 157

Db 152 CWNFTV 157

RESULT 11
US-08-616-241-4
; Sequence 4, Application US/08616241
; Patent No. 5804185
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL RNA EDITING ENZYME
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/616,241
; FILING DATE: Filed Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0239 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 236 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 11777906
US-09-816-241-4

Query Match 19.9%; Score 216.5; DB 1; Length 236;
Best Local Similarity 34.9%; Pred. No. 36-17;
Matches 44; Conservative 31; Mismatches 42; Indels 9; Gaps 4;

QY 35 RDSATSFSLDFGYLR-----NKGQC-HVELLFL-RYISPDWLDLPGRCYRTWTSP 86
DB 33 REACCLYEIKWMSKIRSSGKNTTNHVEVNFIKFTSERDFPSISCSITWFLSWSP 92

QY 87 CYDCARHVADFLRGPNLSLRIFTARLYFCEDRKAEPGRLRLHRAQVQIAIMTKDYFY 146
DB 93 CMECSQAIREFLSRHGGTVLVYVARLFWHMDQ-NRQGLRLVNSGVTIQLMRASEYXH 151

QY 147 CWNTEY 152
DB 152 CWNTEY 157

RESULT 12
US-09-040-482-4
Sequence 4, Application US/09040482
Patent No. 5916556
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Hawkins, Phillip R.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: A NOVEL HUMAN MRNA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,482
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/687,895
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0109 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 236 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1177798
US-09-040-482-4

Query Match 19.9%; Score 216.5; DB 2; Length 236;
Best Local Similarity 34.9%; Pred. No. 36-17;
Matches 44; Conservative 31; Mismatches 42; Indels 9; Gaps 4;

QY 35 RDSATSFSLDFGYLR-----NKGQC-HVELLFL-RYISPDWLDLPGRCYRTWTSP 86
DB 33 REACCLYEIKWMSKIRSSGKNTTNHVEVNFIKFTSERDFPSISCSITWFLSWSP 92

QY 87 CYDCARHVADFLRGPNLSLRIFTARLYFCEDRKAEPGRLRLHRAQVQIAIMTKDYFY 146
DB 93 CMECSQAIREFLSRHGGTVLVYVARLFWHMDQ-NRQGLRLVNSGVTIQLMRASEYXH 151

QY 147 CWNTEY 152
DB 152 CWNTEY 157

RESULT 13
US-09-128-395-4
Sequence 4, Application US/09128395
Patent No. 6087108
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goll, Surya K.
TITLE OF INVENTION: NOVEL RNA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/128,395
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/816,241
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0239 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 236 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 11777906
US-09-128-395-4

Query Match 19.9%; Score 216.5; DB 3; Length 236;

Best Local Similarity 34.9%; Pred. No. 3e-17;
Matches 44; Conservative 31; Mismatches 42; Indels 9; Gaps 4;

QY 35 RRDSSATSFSLDPGYLR-----NKNGC-HVELLFL-RYISDMDLDPGRCYRATWFSWSP 86
Db 33 RKEACLIYEIKWMSKRIWSSGKNTNHEVNFIKFTSERDPFHSISCSITWFLWSNP 92
QY 87 CYDCARHVAADFLKGNPNLSIRIFTARLYFCEDRKAPBGLRLHRAVQIAIMTFKDYFY 146
Db 93 CWFCSQAIRFELSRHPVTLIVYARLFMMDDQ-NRQGLRDLVNSGVITQIMRASEYH 151
QY 147 CWNTEFV 152
Db 152 CWNTEFV 157

RESULT 14
US-08-158-682A-4
Sequence 4, Application US/08158682A
Patent No. 5434058
GENERAL INFORMATION:
APPLICANT: Davidson, Nicholas O.
TITLE OF INVENTION: Apolipoprotein B RNA Editing Protein:
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: 321 No. 5434058th Clark Street, Suite 800
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,682A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Cooley, Ronald B.
REGISTRATION NUMBER: 27,187
REFERENCE/DOCKET NUMBER: ARCD:085
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 744-0090
TELEFAX: (312) 245-4961
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 236 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-158-682A-4

Query Match 19.3%; Score 210; DB 1; Length 236;
Best Local Similarity 35.6%; Pred. No. 1.7e-16;
Matches 48; Conservative 28; Mismatches 43; Indels 16; Gaps 5;

QY 24 RRETYLCYVVK-----RRDSATSFSLDPGYLRNKNGC-HVELLFL-RYISDMDLDPGRCYR 77
Db 33 RKEACLIYEIKWMSKRIWSSGKNTNHEVNFIKFTSERDPFHSISCSITWFLWSNP 83
QY 78 VTWFTSWSPCYDCARHVAADFLKGNPNLSIRIFTARLYFCEDRKAPBGLRLHRAVQIA 137
Db 84 ITWLSWSPCWFCSQAIRFELSRHPVTLIVYARLFMMDDQ-NRQGLRDLVNSGVITQ 142
QY 138 IMTFKDYFYCWNTEFV 152
Db 143 IMRASEYH CWNTEFV 157

RESULT 15
US-08-158-682A-2
Sequence 2, Application US/08158682A
Patent No. 5434058
GENERAL INFORMATION:
APPLICANT: Davidson, Nicholas O.
TITLE OF INVENTION: Apolipoprotein B RNA Editing Protein:
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: 321 No. 5434058th Clark Street, Suite 800
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,682A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Cooley, Ronald B.
REGISTRATION NUMBER: 27,187
REFERENCE/DOCKET NUMBER: ARCD:085
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 744-0090
TELEFAX: (312) 245-4961
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 229 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-158-682A-2

Query Match 18.2%; Score 198; DB 1; Length 229;
Best Local Similarity 32.9%; Pred. No. 4.2e-15;
Matches 56; Conservative 24; Mismatches 62; Indels 28; Gaps 6;

QY 24 RRETYLCYVVK-----RRDSATSFSLDPGYLRNKNGCHVELLFL-RYISDMDLDPGR 74
Db 33 RRETCLEIWMGSHSIWRTSQNT-----NK---HVEVNFIEKFTTERVFCPNT 80
QY 75 CYRTWFTSWSPCYDCARHVAADFLKGNPNLSIRIFTARLYFCEDRKAPBGLRLHRAVQ 134
Db 81 RCSIWFTSWSPCWFCSQAIRFELSRHPVTLIVYARLFMMDDQ-NRQGLRDLVNSGV 139
QY 135 QIAIMTFKDYFYCWNTEFVFNHERTFKAWEGHENSVRSLRRLILPLY 184
Db 140 TIQIMTEGSGY CWNTEFVNSPSNEAHMPPRYPHMRL-----YVLELY 183

Search completed: March 11, 2004, 17:34:57
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comogen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2004, 17:34:20 ; Search time 34 Seconds

(without alignments)
1229.657 Million cell updates/sec

Title: US-09-966-880A-8

Perfect score: 1086
Sequence: 1 MSLLMNRKFLYQFQVW.....ILLFLYVDLDAFTLGL 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppa/PCF_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubppa/PCFUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------------|---------------------|
| 1 | 1086 | 100.0 | 198 | US-09-966-880A-8 | Sequence 8, Appl |
| 2 | 1008 | 92.8 | 198 | US-09-966-880A-2 | Sequence 2, Appl |
| 3 | 390 | 35.9 | 189 | US-10-460-923-5 | Sequence 5, Appl |
| 4 | 390 | 35.9 | 384 | US-09-729-674-174 | Sequence 174, Appl |
| 5 | 390 | 35.9 | 384 | US-10-460-923-2 | Sequence 2, Appl |
| 6 | 369.5 | 34.0 | 222 | US-09-925-300-1639 | Sequence 1639, Appl |
| 7 | 363.5 | 33.5 | 199 | US-10-460-923-7 | Sequence 7, Appl |
| 8 | 349 | 32.1 | 210 | US-10-460-923-4 | Sequence 4, Appl |
| 9 | 300 | 27.6 | 152 | US-10-247-671-159 | Sequence 159, Appl |
| 10 | 238.5 | 22.0 | 195 | US-10-460-923-3 | Sequence 3, Appl |
| 11 | 230 | 21.2 | 219 | US-10-460-923-6 | Sequence 6, Appl |
| 12 | 218.5 | 20.1 | 236 | US-10-157-031-14 | Sequence 14, Appl |
| 13 | 216.5 | 19.9 | 236 | US-10-460-923-8 | Sequence 8, Appl |
| 14 | 211 | 19.4 | 229 | US-09-966-880A-36 | Sequence 36, Appl |
| 15 | 198 | 18.2 | 127 | US-10-104-047-3729 | Sequence 3729, Appl |

| | | | | | | |
|----|------|------|------|----|---------------------|---------------------|
| 16 | 167 | 15.4 | 128 | 15 | US-10-378-029-77 | Sequence 77, Appl |
| 17 | 154 | 14.2 | 151 | 14 | US-10-029-386-34155 | Sequence 34155, A |
| 18 | 84 | 7.7 | 51 | 9 | US-09-864-761-38853 | Sequence 38853, A |
| 19 | 81 | 7.5 | 440 | 13 | US-10-120-319-3 | Sequence 3, Appl |
| 20 | 81 | 7.5 | 440 | 14 | US-10-189-977-3 | Sequence 3, Appl |
| 21 | 81 | 7.5 | 440 | 14 | US-10-392-098-3 | Sequence 3, Appl |
| 22 | 80 | 7.4 | 476 | 9 | US-09-733-524-15 | Sequence 15, Appl |
| 23 | 80 | 7.4 | 476 | 13 | US-10-120-319-5 | Sequence 5, Appl |
| 24 | 80 | 7.4 | 476 | 14 | US-10-189-977-5 | Sequence 5, Appl |
| 25 | 80 | 7.4 | 476 | 14 | US-10-392-098-5 | Sequence 5, Appl |
| 26 | 77.5 | 7.1 | 328 | 15 | US-09-851-873-55 | Sequence 55, Appl |
| 27 | 77.5 | 7.1 | 328 | 15 | US-10-369-493-6748 | Sequence 6748, Appl |
| 28 | 77.5 | 7.1 | 663 | 13 | US-10-080-960-14 | Sequence 14, Appl |
| 29 | 77.5 | 7.1 | 663 | 14 | US-10-247-671-135 | Sequence 135, Appl |
| 30 | 77 | 7.1 | 790 | 14 | US-10-153-668-1164 | Sequence 164, Appl |
| 31 | 75.5 | 7.0 | 382 | 10 | US-09-947-208-25 | Sequence 25, Appl |
| 32 | 75 | 6.9 | 330 | 14 | US-10-265-593-4 | Sequence 4, Appl |
| 33 | 74 | 6.8 | 200 | 10 | US-09-851-873-65 | Sequence 65, Appl |
| 34 | 74 | 6.8 | 707 | 15 | US-10-014-099F-61 | Sequence 61, Appl |
| 35 | 74 | 6.8 | 1291 | 15 | US-10-452-024-122 | Sequence 122, Appl |
| 36 | 73 | 6.7 | 361 | 15 | US-10-439-741-2 | Sequence 2, Appl |
| 37 | 72 | 6.6 | 311 | 16 | US-10-354-437-4 | Sequence 4, Appl |
| 38 | 72 | 6.6 | 357 | 14 | US-10-285-976-35 | Sequence 35, Appl |
| 39 | 72 | 6.6 | 361 | 15 | US-10-312-088-27 | Sequence 27, Appl |
| 40 | 72 | 6.6 | 363 | 14 | US-10-038-288A-4 | Sequence 4, Appl |
| 41 | 72 | 6.6 | 433 | 14 | US-10-038-288A-2 | Sequence 2, Appl |
| 42 | 72 | 6.6 | 704 | 15 | US-10-312-088-26 | Sequence 26, Appl |
| 43 | 72 | 6.6 | 1291 | 15 | US-10-452-024-121 | Sequence 121, Appl |
| 44 | 72 | 6.6 | 1291 | 15 | US-10-452-024-123 | Sequence 123, Appl |
| 45 | 71.5 | 6.6 | 529 | 14 | US-10-053-192-5 | Sequence 5, Appl |

ALIGNMENTS

RESULT 1
US-09-966-880A-8
Sequence 8, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Yasuko
APPLICANT: Matsutani, Masamichi
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966, 880A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 198
TYPE: PRT
ORGANISM: Homo sapiens
US-09-966-880A-8
Query Match 100.0%; Score 1086; DB 9; Length 198;
Best Local Similarity 100.0%; Pred. No. 6.1e-114;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MSLLMNRKFLYQFQVWAKGRRETLGVVVRBDSATSFSDGFLRKNGCHVELL 60
DB 1 MSLLMNRKFLYQFQVWAKGRRETLGVVVRBDSATSFSDGFLRKNGCHVELL 60
OY 61 FLRYISDMDLDPGRRCYRTWTFTSWSPCYCARHVADELRGNPNLSLRFTLRALYFCEDRK 120
DB 61 FLRYISDMDLDPGRRCYRTWTFTSWSPCYCARHVADELRGNPNLSLRFTLRALYFCEDRK 120

QY 121 AEEGRLRLHRAAGVOAIAIMTFKDYFCWNTFVENHRTFKAWEGHLSVRLSRQLRIL 180
 DB 121 AEEGRLRLHRAAGVOAIAIMTFKDYFCWNTFVENHRTFKAWEGHLSVRLSRQLRIL 180
 QY 181 LPLVEVDLRLDAFRTGL 198
 DB 181 LPLVEVDLRLDAFRTGL 198

RESULT 2
 US-09-966-880A-2
 ; Sequence 2, Application US/09966880A
 ; Patent No. US20020164743A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Honjo, Tasuku
 ; APPLICANT: Muramatsu, Masamichi
 ; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
 ; FILE REFERENCE: 06501-088001
 ; CURRENT APPLICATION NUMBER: US/09/966,880A
 ; PRIOR FILING DATE: 2001-09-28
 ; PRIOR APPLICATION NUMBER: PCT/JP00/01918
 ; PRIOR FILING DATE: 2000-03-28
 ; PRIOR APPLICATION NUMBER: JP 11-371382
 ; PRIOR FILING DATE: 1999-12-27
 ; PRIOR APPLICATION NUMBER: JP 11-178999
 ; PRIOR FILING DATE: 1999-06-24
 ; PRIOR APPLICATION NUMBER: JP 11-87192
 ; PRIOR FILING DATE: 1999-03-29
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 198
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-966-880A-2

Query Match 92.8%; Score 1008; DB 9; Length 198;
 Best Local Similarity 92.9%; Pred. No. 3,66-105;
 Matches 183; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
 QY 1 MDSLMMNRKRLYQFNVRWAKGRRETYLCYVVRKRSATSFSLDFGLYLNKNGCHWELL 60
 DB 1 MDSLMMNRKRLYQFNVRWAKGRRETYLCYVVRKRSATSFSLDFGLYLNKNGCHWELL 60
 QY 61 FLRYISDMDLPGRCYRTWTFTSPPCDCAHVADFLRGPNLSLIFRTARLYFCEDRK 120
 DB 61 FLRYISDMDLPGRCYRTWTFTSPPCDCAHVADFLRGPNLSLIFRTARLYFCEDRK 120
 QY 121 AEEGRLRLHRAAGVOAIAIMTFKDYFCWNTFVENHRTFKAWEGHLSVRLSRQLRIL 180
 DB 121 AEEGRLRLHRAAGVOAIAIMTFKDYFCWNTFVENHRTFKAWEGHLSVRLSRQLRIL 180
 QY 181 LPLVEVDLRLDAFRTGL 197
 DB 181 LPLVEVDLRLDAFRTGL 197

RESULT 3
 US-10-460-923-5
 ; Sequence 5, Application US/10460923
 ; Publication No. US2004000951A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MALIM, Michael H.
 ; APPLICANT: SHEEHY, Ann M.
 ; APPLICANT: HARRIS, Reuben S.
 ; APPLICANT: BISHOP, Kate N.
 ; APPLICANT: NEUBERGER, Michael S.
 ; APPLICANT: GADDIS, Nathan C.
 ; APPLICANT: SIMON, James H.M.
 ; TITLE OF INVENTION: DNA Deamination Mediates Innate Immunity to Retroviral Infection
 ; FILE REFERENCE: 22253-74380
 ; CURRENT APPLICATION NUMBER: US/10/460,923

; CURRENT FILING DATE: 2003-06-13
 ; PRIOR APPLICATION NUMBER: US 60/388,513
 ; PRIOR FILING DATE: 2002-06-13
 ; PRIOR APPLICATION NUMBER: US 60/472,952
 ; PRIOR FILING DATE: 2003-05-23
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 189
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-460-923-5

Query Match 35.9%; Score 390; DB 15; Length 189;
 Best Local Similarity 44.9%; Pred. No. 1,26-35;
 Matches 83; Conservative 31; Mismatches 59; Indels 12; Gaps 3;
 QY 6 MNRKRLYQFNVRWAKGRRETYLCYVVRKRSATSFSLDFGLYLNKNGCHWELL 55
 DB 2 MDPPTFTFNPNPVRGRHETLYCYEVRHNDTWLNLQRGFLCNQAPHKHGFLEGR 61
 QY 56 HVELLFLRYISDMDLPGRCYRTWTFTSPPCDCAHVADFLRGPNLSLIFRTARLYF 115
 DB 62 HVELLFLRYISDMDLPGRCYRTWTFTSPPCDCAHVADFLRGPNLSLIFRTARLYF 120
 QY 116 CEDRAEPEGRLRLHRAAGVOAIAIMTFKDYFCWNTFVENHRTFKAWEGHLSVRLSRQ 175
 DB 121 -DDGRCCEGLTTLAAGAKISIMTYSEFKACHMDTFDHQCCPQPMHGDHSDLSGR 179
 QY 176 LRLIL 180
 DB 180 LRLIL 184

RESULT 4
 US-09-729-674-174
 ; Sequence 174, Application US/09729674
 ; Patent No. US20010039335A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: Lavallee, Edward R.
 ; APPLICANT: Collins-Racie, Lisa A.
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Weidberg, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Agostino, Michael J.
 ; APPLICANT: Steindinger II, Robert J.
 ; APPLICANT: Spaulding, Vikki
 ; APPLICANT: Wong, Gordon G.
 ; APPLICANT: Clark, Hilary
 ; APPLICANT: Fecthel, Kim
 ; APPLICANT: Genetics Institute, Inc.
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 ; FILE REFERENCE: 6055-64X
 ; CURRENT APPLICATION NUMBER: US/09/729,674
 ; CURRENT FILING DATE: 2000-12-04
 ; PRIOR APPLICATION NUMBER: 09/539,330
 ; PRIOR FILING DATE: 2000-03-30
 ; NUMBER OF SEQ ID NOS: 263
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 174
 ; LENGTH: 384
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-729-674-174
 Query Match 35.9%; Score 390; DB 9; Length 384;
 Best Local Similarity 44.9%; Pred. No. 2,86-35;
 Matches 83; Conservative 31; Mismatches 59; Indels 12; Gaps 3;
 QY 6 MNRKRLYQFNVRWAKGRRETYLCYVVRKRSATSFSLDFGLYLNKNGCHWELL 55
 DB 2 MDPPTFTFNPNPVRGRHETLYCYEVRHNDTWLNLQRGFLCNQAPHKHGFLEGR 61

Db 197 MDPEPTFFNFNNPFWRGHEHYLCYEVZRMKNDTWLLNQRGLCNQAPKHFLEGR 256
QY 56 HVELLFRLYISDMDLDGRCYRTWTFTSPCYDCARHVADEFKGNPNLSLRIFTARLYE 115
Db 257 HAEELCFDLVTPFWKLDLDQYRTCTFTSPSCQAEAKISKNKXVSLCIFARIY 315
QY 116 CEDKRAPEGLRLHRAAGVQAIAMTFKDYFYCWNTFVENHERTFKAMEGLHENSVALSRQ 175
Db 316 -DDGRCQEGRLTLAAGAKISIMTYSEFKHCWDTFVDHGGCPQPMDDLDHSDLSGR 374
QY 176 LRRIL 180
Db 375 LRAIL 379

RESULT 5
US-10-460-923-2
Sequence 2, Application US/10460923
Publication No. US20040009951A1
GENERAL INFORMATION:
APPLICANT: MALIM, Michael H.
APPLICANT: SHEEHY, Ann M.
APPLICANT: HARRIS, Reuben S.
APPLICANT: BISHOP, Kate N.
APPLICANT: NEUBERGER, Michael S.
APPLICANT: GADDIS, Nathan C.
APPLICANT: SIMON, James H.M.
TITLE OF INVENTION: DNA Deamination Mediates Innate Immunity to Retroviral Infection
FILE REFERENCE: 22253-74380
CURRENT APPLICATION NUMBER: US/10/460,923
CURRENT FILING DATE: 2003-06-13
PRIOR APPLICATION NUMBER: US 60/388,513
PRIOR FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: US 60/472,952
PRIOR FILING DATE: 2003-05-23
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 384
TYPE: PRT
ORGANISM: Homo sapiens
US-10-460-923-2

Query Match 35.9%; Score 390; DB 15; Length 384;
Best Local Similarity 44.9%; Pred. No. 2.8e-35;
Matches 83; Conservative 31; Mismatches 59; Indels 12; Gaps 3;

QY 6 MNRKFLYQFKVWMAKGRRETYLCYVVKRDSATSFSLD--FGYLNRK-----NGC 55
Db 197 MDPEPTFFNFNNPFWRGHEHYLCYEVZRMKNDTWLLNQRGLCNQAPKHFLEGR 256
QY 56 HVELLFRLYISDMDLDGRCYRTWTFTSPCYDCARHVADEFKGNPNLSLRIFTARLYE 115
Db 257 HAEELCFDLVTPFWKLDLDQYRTCTFTSPSCQAEAKISKNKXVSLCIFARIY 315
QY 116 CEDKRAPEGLRLHRAAGVQAIAMTFKDYFYCWNTFVENHERTFKAMEGLHENSVALSRQ 175
Db 316 -DDGRCQEGRLTLAAGAKISIMTYSEFKHCWDTFVDHGGCPQPMDDLDHSDLSGR 374
QY 176 LRRIL 180
Db 375 LRAIL 379

RESULT 6
US-09-925-300-1639
Sequence 1639, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
APPLICANT: Steve Ruden,
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101

CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1639
LENGTH: 222
TYPE: PRT
ORGANISM: Homo sapiens
US-09-925-300-1639

Query Match 34.0%; Score 369.5; DB 9; Length 222;
Best Local Similarity 44.9%; Pred. No. 2.9e-33;
Matches 79; Conservative 24; Mismatches 64; Indels 9; Gaps 4;

QY 11 FLYQFKVWMAKGRRETYLCYVVK--RRDSATSFSLDGYLRN---KNGCHVELLFRLY 64
Db 49 FYQFKVWMAKGRRETYLCYVVK--RRDSATSFSLDGYLRN---KNGCHVELLFRLY 106
QY 65 ISDMDLDGRCYRTWTFTSPCYDCARHVADEFKGNPNLSLRIFTARLYEFCEDKRAPE 124
Db 107 FCDLISPEYTKYQYRTWTFTSPCYDCARHVADEFKGNPNLSLRIFTARLYEFCEDKRAPE 165
QY 125 GLRLHRAAGVQAIAMTFKDYFYCWNTFVENHERTFKAMEGLHENSVALSRQLRIL 180
Db 166 GLRLHRAAGVQAIAMTFKDYFYCWNTFVENHERTFKAMEGLHENSVALSRQLRIL 221

RESULT 7
US-10-460-923-7
Sequence 7, Application US/10460923
Publication No. US20040009951A1
GENERAL INFORMATION:
APPLICANT: MALIM, Michael H.
APPLICANT: SHEEHY, Ann M.
APPLICANT: HARRIS, Reuben S.
APPLICANT: BISHOP, Kate N.
APPLICANT: NEUBERGER, Michael S.
APPLICANT: GADDIS, Nathan C.
APPLICANT: SIMON, James H.M.
TITLE OF INVENTION: DNA Deamination Mediates Innate Immunity to Retroviral Infection
FILE REFERENCE: 22253-74380
CURRENT APPLICATION NUMBER: US/10/460,923
CURRENT FILING DATE: 2003-06-13
PRIOR APPLICATION NUMBER: US 60/388,513
PRIOR FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: US 60/472,952
PRIOR FILING DATE: 2003-05-23
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 199
TYPE: PRT
ORGANISM: Homo sapiens
US-10-460-923-7

Query Match 33.5%; Score 363.5; DB 15; Length 199;
Best Local Similarity 43.5%; Pred. No. 1.2e-32;
Matches 83; Conservative 28; Mismatches 57; Indels 23; Gaps 7;

QY 5 LMRKFLYQFKVWMAKGRRETYLCYVVKRDSATSFSLD--FGYLNRK-----NGC 54
Db 12 LMDPHILFNSFNK---GGRKTYLCYEVZRMKNDTWLLNQRGLCNQAPKHFLEGR 68
QY 55 HVELLFRLYISDMDLDGRCYRTWTFTSPCYDCARHVADEFKGNPNLSLRIFTARLYE 112
Db 69 HAEELCFDLVTPFWKLDLDQYRTCTFTSPSCQAEAKISKNKXVSLCIFARIY 128
QY 113 LKFCEDKRAPE---EGRLHRAAGVQAIAMTFKDYFYCWNTFVENHERTFKAMEGLHENS 169

Db 129 IY-----DYDPLXKALQMLRDAQAQVSIWYDFEKHCMDTFVYDQGCFFQFMDGLDHS 183
QY 170 VRLSRQLRLL 180
Db 184 QALSGRLRAIL 194

RESULT 8
US-10-460-923-4
Sequence 4, Application US/10460923
Publication No. US20040009951A1
GENERAL INFORMATION:
APPLICANT: MALIN, Michael H.
APPLICANT: SHEEHY, Ann M.
APPLICANT: HARRIS, Reuben S.
APPLICANT: BISHOP, Kate N.
APPLICANT: NEUBERGER, Michael S.
APPLICANT: GADDIS, Nathan C.
APPLICANT: SIMON, James H.M.
TITLE OF INVENTION: DNA Deamination Mediates Innate Immunity to Retroviral Infection
FILE REFERENCE: 22253-74380
CURRENT APPLICATION NUMBER: US/10/460,923
CURRENT FILING DATE: 2003-06-13
PRIOR APPLICATION NUMBER: US 60/388,513
PRIOR FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: US 60/472,952
PRIOR FILING DATE: 2003-05-23
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 210
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: mouse orthologue
US-10-460-923-4

Query Match 32.1%; Score 349; DB 15; Length 210;
Best Local Similarity 38.8%; Pred. No. 5.4e-31;
Matches 71; Conservative 36; Mismatches 72; Indels 4; Gaps 2;

QY 5 LMRKKFLYQFKNVWMAKGRRTYLCYVVKRDSATSSLDPGYLRNKGCHVLLFRY 64
Db 20 LLSQTEFKHFHFLRYALDRKDTFLCYEVRKDCSPVLSHHGVFKNDNTHAICFLYW 79
QY 65 ISD---WLDPRGRCRYVMTFTSMSPCYDCARHVADELGNPULSRIFARLYFCEDRK 121
Db 80 FHDKYLKVLSPREERKLTWMSPCFBCAEQVLFATHNLSLDITSSRLYNIRDFEN 139
QY 122 EPEGRRLRHAGVOIAIMTFKDYFCWNTFVENHRTKAMEGLHNSVRLSRQLRLL 181
Db 140 Q-QNLRLVQGAQVAAADLYEFKKCKKFFVNDGRFRPRMKLLTNFRYQDSKLQELLR 198

QY 182 PLY 184
Db 199 PCY 201

RESULT 9
US-10-247-671-159
Sequence 159, Application US/10247671
Publication No. US20030194721A1
GENERAL INFORMATION:
APPLICANT: Wikita, Thomas
APPLICANT: Shiffman, Dov
APPLICANT: Porter, Gordon, J.
APPLICANT: Kaser, Matthew R.
TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
FILE REFERENCE: PA-0050 US
CURRENT APPLICATION NUMBER: US/10/247,671
CURRENT FILING DATE: 2002-09-18
PRIOR APPLICATION NUMBER: 60/323,784
PRIOR FILING DATE: 2001-09-19

NUMBER OF SEQ ID NOS: 186
SOFTWARE: PERL Program
SEQ ID NO 159
LENGTH: 152
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030194721A1 135626CD1
US-10-247-671-159

Query Match 27.6%; Score 300; DB 14; Length 152;
Best Local Similarity 44.9%; Pred. No. 1.2e-25;
Matches 66; Conservative 20; Mismatches 43; Indels 18; Gaps 5;

QY 47 GYLRN--KN-----GCHVEFLFRYISDMWLDPRGRCRYVMTFTSMSPCYD--CARHVA 96
Db 6 GFLHQAKNLGCFYGRHAEIRFLDVLPSLDLDDRAQIRYVMTFTSMSPCFMGCAQGEVRA 65
QY 97 PLRGNPNLSIRIFARLYFCEDRKAP---EGLRLRHAGVOIAIMTFKDYFCWNTFVE 153
Db 66 FLQENTHRLRLIFARLY-----DYDPLXKALQMLRDAQAQVSIWYDFEKHCMDTFVY 120

QY 154 NHERTFKAMEGLHNSVRLSRQLRLL 180
Db 121 RQGCFFQFMDGLDHEHSQLSGRLRAIL 147

RESULT 10
US-10-460-923-3
Sequence 3, Application US/10460923
Publication No. US20040009951A1
GENERAL INFORMATION:
APPLICANT: MALIN, Michael H.
APPLICANT: SHEEHY, Ann M.
APPLICANT: HARRIS, Reuben S.
APPLICANT: BISHOP, Kate N.
APPLICANT: NEUBERGER, Michael S.
APPLICANT: GADDIS, Nathan C.
APPLICANT: SIMON, James H.M.
TITLE OF INVENTION: DNA Deamination Mediates Innate Immunity to Retroviral Infection
FILE REFERENCE: 22253-74380
CURRENT APPLICATION NUMBER: US/10/460,923
CURRENT FILING DATE: 2003-06-13
PRIOR APPLICATION NUMBER: US 60/388,513
PRIOR FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: US 60/472,952
PRIOR FILING DATE: 2003-05-23
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 195
TYPE: PRT
ORGANISM: Homo sapiens
US-10-460-923-3

Query Match 22.0%; Score 238.5; DB 15; Length 195;
Best Local Similarity 34.2%; Pred. No. 1.3e-18;
Matches 63; Conservative 26; Mismatches 84; Indels 11; Gaps 6;

QY 6 MMRKKFLYQFKNVWMAKGRRTYLCYVVKRDSATSSLDPGYLRNKGCHVLLFRY 61
Db 12 MYRDTFSYNFNPIILSRRTVWLCEVKTGSPRP-PLDAKIFRQGVSELYKHBMKF 70
QY 62 LRYISDW-DLDPGRCYVMTFTSMSPCYDCARHVADELGNPULSRIFARLYFCEDRK 120
Db 71 FHWFSKRRKLRHROEYVTVTISMSPTKTRMAFLLEDPKVTLITIFARLYFWMDPD 130
QY 121 APEGRRL--HRAG--VOIAIMTFKDYFCWNTFVENHRTKAMEGLHNSVRLSRQL 176
Db 131 YQ-EALNSLQCKRGRPATWKMIMYDBFQHCWSEFYVSGRELPEPMNMLPKYITLHIML 189
QY 177 RRL 180

DB 190 GEIL 193

RESULT 11
US-10-460-923-6
Sequence 6, Application US/10460923
Publication No. US20040009951A1
GENERAL INFORMATION:
APPLICANT: MALIM, Michael H.
APPLICANT: SHEEHY, Ann M.
APPLICANT: HARRIS, Reuben S.
APPLICANT: BISHOP, Kate N.
APPLICANT: NEUBERGER, Michael S.
APPLICANT: GADDIS, Nathan C.
APPLICANT: SIMON, James H.M.
TITLE OF INVENTION: DNA Deamination Mediates Innate Immunity to Retroviral Infection
FILE REFERENCE: 22253-74380
CURRENT APPLICATION NUMBER: US/10/460,923
CURRENT FILING DATE: 2003-06-13
PRIOR APPLICATION NUMBER: US 60/388,513
PRIOR FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: US 60/472,952
PRIOR FILING DATE: 2003-05-23
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 219
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: mouse orthologue
US-10-460-923-6

Query Match 21.2%; Score 230; DB 15; Length 219;
Best Local Similarity 33.5%; Pred. No. 1,4e-17;
Matches 67; Conservative 35; Mismatches 82; Indels 16; Gaps 8;

QY 5 LNNRRKELYQFNKVR-----WAKGRRETYLCYVVKRSDATSFSLDFGYLRKXNG-CHVE 58
DB 25 LLSSEEFYQFYQRYVHLCYHGMK-PYLQYLEGQFNQAPLK--GCLLSKKQCHAE 80
QY 59 LLEFRISMDLDPGRCYVWTMTSPCYDCARHVADELRSNPULSRIFARLYPCED 118
DB 81 ILFLDKIRSMELSQ--VITTCYLTWSPCPNCAWOLAAKRPDLILHITYSRLLYFWK 137
QY 119 RKAPFGRLRLHAGVOIAIMTFKDYFCMNTFVENHERTFKAMEGLHENSVALSROLR 178
DB 138 RPPQ-KGLCSLWOSGLIVVMDLPQFTDCNTNFV-NPKKPFMPKGLIISRTQRLHR 195
QY 179 ILPLIYVDDLRLDAFRTGL 198
DB 196 I-KESWGLQDLIVDPFNQLQ 214

RESULT 12
US-10-157-031-14
Sequence 14, Application US/10157031
Publication No. US20030108990A1
GENERAL INFORMATION:
APPLICANT: Baranova, A. V.
APPLICANT: Yankovsky, N. K.
APPLICANT: Kozlov, A. P.
APPLICANT: Lobashev, A. V.
APPLICANT: Krukovskaya, L. L.
TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
FILE REFERENCE: 2760-103
CURRENT APPLICATION NUMBER: US/10/157,031
CURRENT FILING DATE: 2002-05-30
NUMBER OF SEQ ID NOS: 415
SOFTWARE: PatentIn version 3.1
SEQ ID NO 14
LENGTH: 236

TYPE: PRT
ORGANISM: Homo sapiens
US-10-157-031-14

Query Match 20.1%; Score 218.5; DB 14; Length 236;
Best Local Similarity 34.9%; Pred. No. 3e-16;
Matches 44; Conservative 31; Mismatches 42; Indels 9; Gaps 4;

QY 35 RDSATSFSLDFGYLR-----NKNGC-HVELLFL-RYISDMDLDPGRCYVWTMTSP 86
DB 33 RKACILYEIKWMSKIRSSGKNTTNHVEVAFIKFTSERDFHSMSCSIWFASMP 92
QY 87 CYDCARHVADELRSNPULSRIFARLYPCEDRKAPEGLRLHAGVOIAIMTFDYFY 146
DB 93 CWECSCQAIREFLSRHGVTLVIVARLFWMDDQ-NRQGLRDLVNSGVTIQIMRASEYH 151
QY 147 CWNTFV 152
DB 152 CWNTFV 157

RESULT 13
US-10-460-923-8
Sequence 8, Application US/10460923
Publication No. US20040009951A1
GENERAL INFORMATION:
APPLICANT: MALIM, Michael H.
APPLICANT: SHEEHY, Ann M.
APPLICANT: HARRIS, Reuben S.
APPLICANT: BISHOP, Kate N.
APPLICANT: NEUBERGER, Michael S.
APPLICANT: GADDIS, Nathan C.
APPLICANT: SIMON, James H.M.
TITLE OF INVENTION: DNA Deamination Mediates Innate Immunity to Retroviral Infection
FILE REFERENCE: 22253-74380
CURRENT APPLICATION NUMBER: US/10/460,923
CURRENT FILING DATE: 2003-06-13
PRIOR APPLICATION NUMBER: US 60/388,513
PRIOR FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: US 60/472,952
PRIOR FILING DATE: 2003-05-23
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8
LENGTH: 236
TYPE: PRT
ORGANISM: Homo sapiens
US-10-460-923-8

Query Match 19.9%; Score 216.5; DB 15; Length 236;
Best Local Similarity 34.9%; Pred. No. 5.1e-16;
Matches 44; Conservative 31; Mismatches 42; Indels 9; Gaps 4;

QY 35 RDSATSFSLDFGYLR-----NKNGC-HVELLFL-RYISDMDLDPGRCYVWTMTSP 86
DB 33 RKACILYEIKWMSKIRSSGKNTTNHVEVAFIKFTSERDFHSMSCSIWFASMP 92
QY 87 CYDCARHVADELRSNPULSRIFARLYPCEDRKAPEGLRLHAGVOIAIMTFDYFY 146
DB 93 CWECSCQAIREFLSRHGVTLVIVARLFWMDDQ-NRQGLRDLVNSGVTIQIMRASEYH 151
QY 147 CWNTFV 152
DB 152 CWNTFV 157

RESULT 14
US-09-966-880A-36
Sequence 36, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Muramatsu, Masamichi

Mon Mar 15 09:44:26 2004

us-09-966-880a-8.rapb

Page 6

Search completed: March 11, 2004, 17:33:27
Job time : 35 secs

TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966,880A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 36
LENGTH: 229
TYPE: PRT
ORGANISM: Mus musculus
US-09-966-880A-36

Query Match 19.4%; Score 211; DB 9; Length 229;
Best Local Similarity 37.1%; Pred. No. 2e-15;
Matches 49; Conservative 24; Mismatches 49; Indels 10; Gaps 4;

QY 24 RRETYLCYVVK--RRDSATSFSLDFGLRNKNGCHVELLFL-RYISDMDLDPGRCYVTW 80
DB 33 RKETCLYEINWGRSHV-----WRHTSGNTSNHVEVNFLEFTEYRFRNTSCITW 86

QY 81 FTSWSPCYDCARHVAADFLRGNPNLSLRTFARLYFCEDRKABEGGLRLHRAQVOIAIMT 140
DB 87 FLSWSPGECRSALTEFLSRHPYVTLFIYARLYHHTDOR-NRQGLRDLISSGVTIQIMT 145

QY 141 FKDYFCMNTFV 152
DB 146 EQEYCYCMRNV 157

RESULT 15

US-10-104-047-3729
Sequence 3729, Application US/10104047
Publication No. US20030236392A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: NO. US20030236392A1e1 full length cDNA
FILE REFERENCE: HL-A0105
CURRENT APPLICATION NUMBER: US/10/104,047
CURRENT FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER:
PRIOR FILING DATE:
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3729
LENGTH: 127
TYPE: PRT
ORGANISM: Homo sapiens
US-10-104-047-3729

Query Match 18.2%; Score 198; DB 15; Length 127;
Best Local Similarity 36.8%; Pred. No. 2.8e-14;
Matches 46; Conservative 19; Mismatches 54; Indels 6; Gaps 4;

QY 61 FLRYISDM-DLDPGRCYVTWFTWSWSPCYDCARHVAADFLRGNPNLSLRTFARLYFCEDR 119
DB 3 FFHMFSSKRRKLRHQSEYEVWYISWSPCTKTRDMATFLAEDPKVTLTIFFVRLYYFMDP 62

QY 120 KABEGGLRL--HRAQ--VOIAIMTFDYCYCNTFVENHRTFKAMEGGLHNSVRLSRQ 175
DB 63 DYQ-EALRSLCQRRDGPRAIMKIMYDFQHCWSKRVYSQRELFEFNNLFPKYILLHIM 121

QY 176 LRRIL 180
DB 122 LGEIL 126

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 13, 2004, 09:16:37 ; Search time 66 Seconds
(without alignments)
1664.853 Million cell updates/sec

Title: US-09-966-880A-8
Perfect score: 1086
Sequence: 1 MDSLLMNRKRLYQFKVNRW.....ILPLVEVDLDRDAFRLGL 198

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=x1h
-O=/cg2_1/USPTO.spool/US09966880/runat_11032004_172742_26608/app_query.fasta_1.391
-DB=Issued_Patents_NA -QPWT=fastap -SUFFIX=nti -MINMATCH=0.1 -LOOPEC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdt
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09966880 @CGN 1.1 44 @runat_11032004_172742_26608 -NCU=6 -ICPU=3
-NO WMAP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGIOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cg2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cg2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cg2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cg2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cg2_6/ptodata/2/ina/6C.COMB.seq:*
6: /cg2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|---------|-------|---------------------------------------|
| 1 | 369.5 | 34.0 | 610 | 1 | US-08-816-241-2 Sequence 2, Appli |
| 2 | 369.5 | 34.0 | 610 | 3 | US-09-128-395-2 Sequence 2, Appli |
| 3 | 271 | 25.0 | 891 | 1 | US-08-687-895-2 Sequence 2, Appli |
| 4 | 271 | 25.0 | 891 | 2 | US-09-040-482-2 Sequence 2, Appli |
| 5 | 210 | 19.3 | 879 | 1 | US-08-158-682A-3 Sequence 3, Appli |
| 6 | 201 | 18.5 | 879 | 1 | US-08-158-682A-1 Sequence 1, Appli |
| 7 | 201 | 18.5 | 879 | 1 | US-08-015-203-1 Sequence 1, Appli |
| 8 | 87.5 | 8.1 | 1806 | 3 | US-09-068-655-10 Sequence 10, Appli |
| 9 | 87.5 | 8.1 | 4713 | 3 | US-09-068-655-3 Sequence 3, Appli |
| 10 | 86 | 7.9 | 1230025 | 4 | US-09-198-452A-1 Sequence 1, Appli |
| 11 | 80.5 | 7.4 | 1356 | 4 | US-09-107-532A-2790 Sequence 2790, Ap |
| 12 | 80 | 7.4 | 7077 | 4 | US-09-221-017B-123 Sequence 123, App |

| | | | | | | |
|------|------|-----|--------|---|---------------------|--------------------|
| C 13 | 79.5 | 7.3 | 591 | 4 | US-09-252-991A-7743 | Sequence 7743, Ap |
| 14 | 79.5 | 7.3 | 816 | 4 | US-09-252-991A-7748 | Sequence 7748, Ap |
| C 15 | 78 | 7.2 | 1053 | 4 | US-09-252-991A-6143 | Sequence 6143, Ap |
| 16 | 78 | 7.2 | 1167 | 4 | US-09-252-991A-6408 | Sequence 6408, Ap |
| 17 | 77.5 | 7.1 | 1209 | 6 | 5352575-4 | Patent No. 5352575 |
| 18 | 77.5 | 7.1 | 4446 | 4 | US-09-976-594-668 | Sequence 668, App |
| 19 | 76.5 | 7.0 | 1905 | 4 | US-09-252-991A-7829 | Sequence 7829, Ap |
| C 20 | 75.5 | 7.0 | 3043 | 4 | US-09-049-698-16 | Sequence 16, Appl |
| C 21 | 75.5 | 7.0 | 3181 | 4 | US-09-049-698-18 | Sequence 18, Appl |
| C 22 | 75 | 6.9 | 323 | 4 | US-09-210-952-16 | Sequence 16, Appl |
| C 23 | 74 | 6.8 | 269223 | 4 | US-09-596-002-41 | Sequence 41, Appl |
| 24 | 73.5 | 6.8 | 852 | 4 | US-09-489-035A-467 | Sequence 467, App |
| 25 | 73.5 | 6.8 | 2629 | 1 | US-08-200-807-1 | Sequence 1, Appli |
| 26 | 73.5 | 6.8 | 2629 | 1 | US-08-488-305A-1 | Sequence 1, Appli |
| 27 | 73 | 6.7 | 801 | 4 | US-09-252-991A-2076 | Sequence 2076, Ap |
| 28 | 73 | 6.7 | 909 | 4 | US-09-252-991A-2147 | Sequence 2147, Ap |
| C 29 | 72.5 | 6.7 | 1578 | 4 | US-09-107-532A-2089 | Sequence 2089, Ap |
| C 30 | 72.5 | 6.7 | 9179 | 4 | US-09-453-702B-162 | Sequence 162, App |
| C 31 | 72.5 | 6.7 | 13188 | 4 | US-08-961-527-70 | Sequence 70, Appl |
| C 32 | 72 | 6.6 | 531 | 4 | US-09-252-991A-4510 | Sequence 4510, Ap |
| 33 | 72 | 6.6 | 1464 | 4 | US-09-489-035A-2639 | Sequence 2639, Ap |
| C 34 | 72 | 6.6 | 1672 | 4 | US-09-495-066-1 | Sequence 1, Appli |
| 35 | 72 | 6.6 | 4201 | 3 | US-08-080-255-4 | Sequence 4, Appli |
| 36 | 72 | 6.6 | 4201 | 3 | US-08-465-713-4 | Sequence 4, Appli |
| 37 | 72 | 6.6 | 4201 | 5 | PCT-US93-05857-4 | Sequence 4, Appli |
| 38 | 72 | 6.6 | 5751 | 4 | US-09-023-655-1415 | Sequence 1415, Ap |
| 39 | 72 | 6.6 | 11907 | 3 | US-08-061-376-4 | Sequence 4, Appli |
| 40 | 72 | 6.6 | 14255 | 1 | US-08-320-559-1 | Sequence 1, Appli |
| 41 | 72 | 6.6 | 14255 | 1 | US-08-327-392-1 | Sequence 1, Appli |
| 42 | 72 | 6.6 | 14255 | 1 | US-08-306-691B-55 | Sequence 55, Appl |
| 43 | 72 | 6.6 | 14255 | 3 | US-08-545-860D-1 | Sequence 1, Appli |
| 44 | 72 | 6.6 | 14255 | 5 | PCT-US84-0446-1 | Sequence 1, Appli |
| C 45 | 72 | 6.6 | 536165 | 4 | US-09-214-808-1 | Sequence 1, Appli |

ALIGNMENTS

RESULT 1
US-08-816-241-2
Sequence 2, Application US/08816241
Patent No. 5804185
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goll, Surya K.
TITLE OF INVENTION: NOVEL RNA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/816,241
FILING DATE: Filed Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0239 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:

```

SEQUENCE CHARACTERISTICS::
LENGTH: 610 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSTUT09
CLONE: 1646833

```


DB 520 TTTAAACCACTTTCGACTTCTGAAAAAGAGGCTACGGGAGACTCTC 567

RESULT 3

US-08-687-895-2
Sequence 2, Application US/08687895
Patent No. 5747319

GENERAL INFORMATION:

APPLICANT: Au-Young, Janice
APPLICANT: Hawkins, Phillip R.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: A NOVEL HUMAN RNA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,895

FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0109 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-4166

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 891 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

IMMEDIATE SOURCE:

LIBRARY: MUSCNOT1

CLONE: 57953

US-08-687-895-2

Alignment Scores:

Pred. No.: 2,92e-25 Length: 891
Score: 271.00 Matches: 63
Percent Similarity: 50.00% Conservative: 25
Best Local Similarity: 35.80% Mismatches: 76
Query Match: 24.95% Indels: 12
DB: 1 Gaps: 5

US-09-966-880A-8 (1-198) x US-08-687-895-2 (1-891)

QY 11 PheLeuTyrglnPheLeuAsnValArgTrpAlaLysGlyArgArgGlnThrTyrlencys 30
DB 306 TTTAAATTCACGATTCGGATGTGAGTACACTTCGGAGAACACGCTTCCTCTGC 365
QY 31 TyrValValLysArgAspSerAlaThrSerPheSerLeuAspPheGlyTyrlencys 50
DB 366 TATGTGTGAGGACAGGGGCAAGGGGGCCAGATGCGAGGCACTTCGGGATACCTAAG 425
QY 51 AsnLysAsn---GlyCysHisValGlnLeuLeuPheLeuArgTyrIleSerAspTrpAsp 69
DB 426 GATGACATCGCGGCTGCCATGCAGAGAGCTTCTTCAACACCATCCTGCA---GCC 482
QY 70 LeuAspProGlyArgCysTyArgValThrTrpPheThrSerTrpSerProCysTyArgp 89
DB 483 TTCGACCCGACCTCGGCTACATGTCACCTGTGTGTCTCTCCAGCCCTGTGCAAG 542
QY 90 CysAlaArgHisValAlaAspPheLeuArgGlyAsnProAsnLeuSerLeuArgIlePhe 109

DB 543 TGTGTGACCGGATATAYCAAAACCTTGACAGACCAAGAACCTGCGTCTGCTATCTG 602

QY 110 ThrAlaArgLeuTyrgPheCysGlnAspArgLysAlaGlnProGlu-----GlyLeu 126

DB 603 GTGGGTGACCTCTCTGTGGAG-----GAGCCGAGATCCAGGCTGCTGTG 650

QY 127 ArgArgLeuHisArgAlaGlyValGlnIleAlaIleMetThrPheLysAspTyrPheTyr 146

DB 651 AAGAGCTAAGAGAGGCTGGTGTAACTGCGCATCATGAGCCCAAGACTTCAATAT 710

QY 147 CysTrpAsnThrPheValGlnAsnHisGlu-----ArgThrPheLysAlaTrpGlu 163

DB 711 GTCTGCAGAAATTTCTGTGACAGAGAGGCTGAATCCAAAGCCTTTCARCCCTGGAG 770

QY 164 GlyLeuHisGlnAsnSerValArgLeuSerArgGlnLeuArgIle 179

DB 771 GACATTCAGAGAACTTCTATCTACGAGAGAAATTGGCAGACATC 818

RESULT 4

US-09-040-482-2
Sequence 2, Application US/09040482
Patent No. 5916556

GENERAL INFORMATION:

APPLICANT: Au-Young, Janice
APPLICANT: Hawkins, Phillip R.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: A NOVEL HUMAN RNA EDITING ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,482

FILING DATE:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/687,895

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0109 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 891 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

IMMEDIATE SOURCE:

LIBRARY: MUSCNOT1

CLONE: 57953

US-09-040-482-2

Alignment Scores:

Pred. No.: 2,92e-25 Length: 891
Score: 271.00 Matches: 63
Percent Similarity: 50.00% Conservative: 25
Best Local Similarity: 35.80% Mismatches: 76
Query Match: 24.95% Indels: 12
DB: 2 Gaps: 5

41

CLASSIFICATION: 435

US-09-966-880A-8 (1-198) x US-08-158-682A-3 (1-879

QY 175 GlnleuargargileuLeuProleuTyr 184
Db 572 -----TAGCTACTGGAACCTCTAC 589

RESULT 8

US-09-068-655-10/c
; Sequence 10, Application US/09068655A
; Patent No. 6136579
; GENERAL INFORMATION:
; APPLICANT: JACKSON, Ronald James
; TITLE OF INVENTION: METHOD OF PRODUCING ALPHA 2,3-SIALYLTRANSFERASE
; FILE REFERENCE: 026579-186
; CURRENT APPLICATION NUMBER: US/09/068,655A
; EARLIER FILING DATE: 1998-08-31
; EARLIER APPLICATION NUMBER: AU PN 6587
; EARLIER FILING DATE: 1995-11-15
; EARLIER APPLICATION NUMBER: PCT/AU96/00725
; EARLIER FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1806
; TYPE: DNA
; ORGANISM: Myxoma virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(439)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1405)..(1806)
US-09-068-655-10

Alignment Scores:

| | | | |
|------------------------|--------|---------------|------|
| Pred. No.: | 0.339 | Length: | 1806 |
| Score: | 87.50 | Matches: | 38 |
| Percent Similarity: | 41.07% | Conservative: | 8 |
| Best Local Similarity: | 33.93% | Mismatches: | 43 |
| Query Match: | 8.06% | Indels: | 23 |
| DB: | 3 | Gaps: | 7 |

US-09-966-880A-8 (1-198) x US-09-068-655-10 (1-1806)

QY 20 TPpAlaYsGlyArGArGluThrTyrlEuCyStYrValVAllySArGArGAspSerAla 35
Db 888 TCGTTCCTGCTCAAGCCGCCGACTATATCGCTATATACATGCTTACGAACAACGAC 829
QY 40 ThrsrPheserleuaspPheGlyTyrlEuArGAnlySArGlyCyShIeValGluLeu 59
Db 828 GCGTCGACCTGCTAGATTGGAAGAAACCGCTCGACGTCGGCGCGCTCGGCGT 769
QY 60 leupheleuArGtyrlIeserasp-----TriAspleuasp 71
Db 768 TCGTACGAATCTCAATCCGAATATACGTACGATCGCGCTGCGTATACAGACACGT 709
QY 72 ProGlyArGArGtyrArGValThrTrpPheThrsrIser-----ProCyStYr 88
Db 708 CCGGAACGTTGACGCTTCTTACGATG-----GAATGTCGCGCTCGTGAACCGGCTTAC 655
QY 89 AspCyAlaArGhIeValAlaaspPheleuArG-----GlyAsnPro----- 102
Db 654 ACCTGTCGACG-----CGTACAGATTACGGGATTCGGGTACCCCAATGCGGCGCTA 601
QY 103 ---AsnleuSerleuArGlyIepheThrAlaArGleu 113
Db 600 ACCAATTATCCATT---ATTACACGCGCTATACTA 568

RESULT 9

US-09-068-655-3/c
; Sequence 3, Application US/09068655A
; Patent No. 6136579
; GENERAL INFORMATION:
; APPLICANT: JACKSON, Ronald James

; TITLE OF INVENTION: METHOD OF PRODUCING ALPHA 2,3-SIALYLTRANSFERASE

; FILE REFERENCE: 026579-186
; CURRENT APPLICATION NUMBER: US/09/068,655A
; EARLIER FILING DATE: 1998-08-31
; EARLIER APPLICATION NUMBER: AU PN 6587
; EARLIER FILING DATE: 1995-11-15
; EARLIER APPLICATION NUMBER: PCT/AU96/00725
; EARLIER FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3

; LENGTH: 4713
; TYPE: DNA
; ORGANISM: Myxoma virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(549)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1515)..(2081)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2084)..(3745)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3776)..(4434)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4440)..(4712)
US-09-068-655-3

Alignment Scores:

| | | | |
|------------------------|--------|---------------|------|
| Pred. No.: | 1.44 | Length: | 4713 |
| Score: | 87.50 | Matches: | 38 |
| Percent Similarity: | 41.07% | Conservative: | 8 |
| Best Local Similarity: | 33.93% | Mismatches: | 43 |
| Query Match: | 8.06% | Indels: | 23 |
| DB: | 3 | Gaps: | 7 |

US-09-966-880A-8 (1-198) x US-09-068-655-3 (1-4713)

QY 20 TPpAlaYsGlyArGArGluThrTyrlEuCyStYrValVAllySArGArGAspSerAla 35
Db 998 TCGTTCCTGCTCAAGCCGCCGACTATATCGCTATATACATGCTTACGAACAACGAC 939
QY 40 ThrsrPheserleuaspPheGlyTyrlEuArGAnlySArGlyCyShIeValGluLeu 59
Db 938 GCGTCGACCTGCTAGATTGGAAGAAACCGCTCGACGTCGGCGCGCTCGGCGT 879
QY 60 leupheleuArGtyrlIeserasp-----TriAspleuasp 71
Db 878 TCGTACGAATCTCAATCCGAATATACGTACGATCGCGCTGCGTATACAGACACGT 819
QY 72 ProGlyArGArGtyrArGValThrTrpPheThrsrIser-----ProCyStYr 88
Db 818 CCGGAACGTTGACGCTTCTTACGATG-----GAATGTCGCGCTCGTGAACCGGCTTAC 765
QY 89 AspCyAlaArGhIeValAlaaspPheleuArG-----GlyAsnPro----- 102
Db 764 ACCTGTCGACG-----CGTACAGATTACGGGATTCGGGTACCCCAATGCGGCGCTA 711
QY 103 ---AsnleuSerleuArGlyIepheThrAlaArGleu 113
Db 710 ACCAATTATCCATT---ATTACACGCGCTATACTA 678

RESULT 10

US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment thereof and uses thereof, in particular for the diagnosis, pre-


```

LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature

```

Alignment Scores:

```

Pred. No.: 9.63e+03 Length: 1230025
Score: 86.00 Matches: 55
Percent Similarity: 36.89% Conservative: 21
Best Local Similarity: 26.70% Mismatches: 57
Query Match: 7.92% Indels: 73
Gaps: 10

```

US-09-966-880A-8 (1-198) x US-09-198-452A-1 (1-1230025)

```

QY 9 ArglySphE---LeuTyGlnPheIysAsnValArgTPaIaIySgIyArgArgIuThr 27
DB 289205 AGAAATATCTCTTATATCAAGAACCCACATGAGTCAATTAATTCCTACAGAAATCAG 289146
QY 28 TyIleuCyStyVal----- 32
DB 289145 ATCCCTTATACATCGTTGAGGGCGGTGAGCCGCCGCTCCACATCTACATCA 289086
QY 33 -----ValIyArgArgAspSerAlaThrSerPheSerLeuAspPheGlyTyIleuArg 50
DB 289085 AGCCTGTAAAGATTATTAACATCAATCACTTAATTTCTTGTTGCTTGCTTAATTA--- 289029
QY 51 AsnIyAsnGlyCyShiValIuIleuPheIuArgTyIle-SerAspTPaAspLe 70

```

```

DB 289028 -----TTACGATATCTTACAGACCTTATATG 289002
QY 70 uAsPProGlyArGyCyStyArGyValInrTTPheInrSerTProCyStyAsp-- 89
DB 289001 TAGGCTCT-----TGTTTTGCTCTATTTTATTAATCTCGAGCTCTGCTGATGA 288948
QY 90 -----CysAlaArgShiValAlaAspPheIuArgIyAsnDr 102
DB 288947 CCGCAGCATATGATGATGCTGCTGCTCAGAGAG-----AACAGGATCC 288903
QY 102 AsnIuSerIuArgIlePheThraIaArgIeuTyPheCySgIuAspArgIyAlaG1 122
DB 288902 TCCTCTGTAATAGGCTTTTACTTGC-----TTCTGTATAGGCTGACAGAAA 288852
QY 122 uPro-GluGlyIuArgIleuHsArgIlaGlyValGlnIleAlaIle-MetTrPhe 141
DB 288851 GCCAAGAGGTGTCATCATATGATCCGCACTATTCGTTGTATAGTGAATACA--- 288795
QY 142 LysAspTyPheTyCyStyTPaInrPheValGluAsnHsGluArgThrPheIyAla 161
DB 288794 ---CATATTCAGATGC---CTTACTTCTTATCACTCCAGAA----- 288756
QY 162 TrpGluGlyIuHsGluAsnSerValArgIuSerArgIleuArgIleu 181
DB 288755 -----CTAAAAAAGTCTTACTC 288738
QY 182 ProIeuTyArgIu 185
DB 288737 AGTCTATACGAA 288726

RESULT 11
US-09-107-532A-2790
Sequence 2790, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESS: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 2790:
SEQUENCE CHARACTERISTICS:
LENGTH: 1356 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)

```

```

?      HYPOTHEITICAL: NO
?      ANTI-SENSE: NO
?      ORIGINAL SOURCE:
?      ORGANISM: Enterococcus faecium
?      FEATURE:
?      NAME/KEY: misc feature
?      LOCATION: (B) LOCATION 1..1356
?      SEQUENCE DESCRIPTION: SEQ ID NO: 2790
IS-09-107-532A-2790

```

| | |
|------------------------|--------|
| Alignment Scores: | |
| Pred. No.: | 1.75 |
| Score: | 80.50 |
| Percent Similarity: | 37.97% |
| Best Local Similarity: | 24.68% |
| Query Match: | 7.41% |
| DB: | 4 |
| Length: | 1366 |
| Matches: | 39 |
| Conservative: | 21 |
| Mismatches: | 59 |
| Indels: | 39 |
| Gaps: | 8 |

US-09-966-880A-8 (1-198) x US-09-107-532A-2790 (1-1356)

QY 45 LeuDbgSnlrYkSnglYcYshSvalGluLeuLeuPheLeuAsGTYrLlSerAspTrp 68
 Db 433 CTGATTACGAAACGGCGCAACGCTTAAAGAAATCATTTTACGCTACGCAAGAAAAATGG 432
 QY 69 AspLeuAspProGluYarGyrCysYrArGValThrTrpPheThrSerTrpSerProCysYr 88
 Db 493 GATTACACCTGCT-----TTGTTGAGTGAGTTAGAAAAAGAA 531
 QY 89 Asp-----CysAlaArgHisValAlaAspPheLeuArgGluAsnProAsnLeuSerLeu 106
 Db 532 AATGTTCCTGGTGTAGTTAGTTAGTCGCATGCTCCAGGTTATCTCGCGAAGATGCA 531
 QY 107 ArgIlePheThrAlaArgLeuTYrPheCysGluAspArg-----LysAlaGluPro 123
 Db 592 GGTCTTTTCATCAACGCGCATACACT---GAAGATCAATTGATGCTGAAGCAAGAACCA 648
 QY 124 -----GluGlyLeuArgArgLeuHis-----Arg 131
 Db 649 TTCATGCTTTGGGTGATCGAAGCGGCTTAAAGATTTGAGAAATCATCATCATCAACAG 708
 QY 132 AlaIleValGlnIleAlaIleMetThrPheLysAspTYrPheTYrCysTrpAsnThrPhe 151
 Db 709 GCTGGCTTGAATGTCATCATCACAA-----GATGATATG 741
 QY 152 ValGluAsnHisGluArgThrPheLysAlaTrpGluGlyLeuHisGluAsnSerValArg 171
 Db 742 ACTCCTTATGCTGACGTAAAGTACGATTCGTTAGTAAAGCCGATACAGCGATGTCCT 801
 QY 172 LeuSerArgGlnLeuArgArgIleLeuLeuProLeuTYrGluValAspAspLeu 169
 Db 802 CTTCCT-----TTGTGGCGGAATTTGAAACTGTGAGAAAGATA 840

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PPI182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PPI146
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Montoy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-020021.000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 123:
SEQUENCE CHARACTERISTICS:
LENGTH: 7077 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...7077
US-09-221-017B-123

```

| | | |
|------------------------|--------|-----------------|
| Alignment Scores: | | |
| Pred. No.: | 24.4 | 7077 |
| Score: | 80.00 | Matches: 22 |
| Percent Similarity: | 47.65% | Conservative: 8 |
| Best Local Similarity: | 74.92% | Mismatches: 17 |
| Query Match: | 7.37% | Indels: 16 |
| DB: | 4 | Gaps: 3 |

US-03-966-880A-8 (1-198) x US-03-221-017B-123 (1-7077)

| | | | | |
|----|--|------|--|------|
| Oy | | 4 | LeuLeuMetAsnArgTgATgLySPhelLeuTyGlnPheLySAsnValAlaGTTPAlALyGly | 23 |
| | | | ::::: ::: ::: ::: ::: | |
| Db | | 3900 | TTCATATATAAATCCAAAACGTTTATATATATAGATCGTCACCGCTGCATTGGGAAAAAAC | 3959 |
| Oy | | 24 | ArgArgGluThrTYrLeuCySTyrValValLys-----ArgArgAspSerAlaThr | 40 |
| | | | ::: ::: ::: ::: | |
| Db | | 3960 | GCCGAGAACAATTTCCTTGAAAGCATATAAAAAGCAAGCACCCGCCGAAGTTCCTG--- | 4016 |
| Oy | | 41 | SerPheSerLeuAspPheGlyTYrLeuArgAsnLySAnGlyCySHisValGluLeuLeu | 60 |
| | | | ::: ::: ::: ::: | |
| Db | | 4017 | -----CGAAGCTCCGAGCGGGTGC-----CTTGCT | 4040 |
| Oy | | 61 | PheLeuArg | 63 |
| | | | | |
| Db | | 4041 | TTTTGTGCT | 4049 |

```

RESULT 13
US-09-252-991A-7743/c
; Sequence 7743, Application US/0925291A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

```


| | | | |
|----|-----|--|-----|
| Db | 413 | CGATGGCAGCCGTTACTGATCTGCAACGGCTACGACGGCAATCCCGGC | 354 |
| Qy | 178 | GILEuleuProleuTyrgluValAspAspleu---ArgAepAlaPheargThrleuGI | 197 |
| Db | 353 | AGGCCAACTACCGGGCGATCCAGTACGACACCTTATGCGTATGCTGCCCAAGCCGAGG | 294 |
| Qy | 197 | Y 197 | |
| Db | 293 | C 293 | |

Search completed: March 13, 2004, 14:53:23
 Job time : 328 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 13, 2004, 10:11:37 ; Search time 275 Seconds

(Without alignments)
2650.750 Million cell updates/sec

Title: US-09-966-880A-8

Profile score: 1086

Sequence: 1 MDSLMNRKFLYQFKYRW.....ILLPLYVDLRLDAFRLGL 198

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n model -DEV=xlh
-DB=PubMed/US09966880/runat_11032004_172743_26651/app_query.fasta_1.391
-LOOPEL=0 -LOOPEXT=0 -UNITS=DIC -START=1 -END=1 -MATRIX=D1osum62
-TRAN=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09966880 @CGN 1.1.164 @runat_11032004_172743_26651
-NCPU=6 -ICPT=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLICED=100
-LONGLOG -DEV-TIMEOUT=120 -WARN-TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELXT=7

Database: Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10E_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result | Score | Query | Match | Length | ID | Description |
|--------|-------|-------|-------|--------|----|-------------|
|--------|-------|-------|-------|--------|----|-------------|

| | | | | | | |
|----|-------|-------|--------|----|----------------------|--------------------|
| 1 | 1086 | 100.0 | 2818 | 9 | US-09-966-880A-7 | Sequence 7, Appl |
| 2 | 1008 | 92.8 | 2440 | 9 | US-09-966-880A-1 | Sequence 10, Appl |
| 3 | 644.5 | 59.3 | 6564 | 9 | US-09-966-880A-10 | Sequence 1, Appl |
| 4 | 644.5 | 59.3 | 11204 | 9 | US-09-966-880A-35 | Sequence 35, Appl |
| 5 | 505 | 46.5 | 271 | 9 | US-09-966-880A-13 | Sequence 13, Appl |
| 6 | 390 | 35.9 | 1155 | 15 | US-10-460-923-1 | Sequence 1, Appl |
| 7 | 390 | 35.9 | 1534 | 9 | US-09-729-674-173 | Sequence 173, App |
| 8 | 369.5 | 34.0 | 987 | 9 | US-09-925-300-699 | Sequence 699, App |
| 9 | 369.5 | 34.0 | 1143 | 13 | US-10-098-841-268 | Sequence 268, App |
| 10 | 363.5 | 33.5 | 1348 | 9 | US-09-880-107-3282 | Sequence 3282, App |
| 11 | 348 | 32.0 | 883 | 14 | US-10-247-671-62 | Sequence 62, Appl |
| 12 | 343 | 31.6 | 1008 | 15 | US-10-131-827-8890 | Sequence 8890, App |
| 13 | 308.5 | 28.4 | 2151 | 9 | US-09-822-830A-359 | Sequence 359, App |
| 14 | 262 | 24.1 | 201143 | 12 | US-10-240-425-1099 | Sequence 1099, App |
| 15 | 260 | 23.5 | 148 | 9 | US-09-966-880A-12 | Sequence 12, Appl |
| 16 | 257 | 23.7 | 371 | 10 | US-09-918-995-30237 | Sequence 30237, A |
| 17 | 249.5 | 23.0 | 2913 | 15 | US-10-027-632-112023 | Sequence 112023, |
| 18 | 249.5 | 22.0 | 2913 | 15 | US-10-027-632-112024 | Sequence 112024, |
| 19 | 238.5 | 20.1 | 1404 | 14 | US-10-198-846-13472 | Sequence 13472, A |
| 20 | 218.5 | 20.1 | 879 | 14 | US-10-157-031-13 | Sequence 13, Appl |
| 21 | 212 | 19.5 | 116 | 9 | US-09-866-880A-14 | Sequence 14, Appl |
| 22 | 209 | 19.2 | 2773 | 15 | US-10-104-047-1759 | Sequence 1759, App |
| 23 | 206 | 19.0 | 390 | 9 | US-09-796-692-9051 | Sequence 9051, App |
| 24 | 206 | 19.0 | 390 | 14 | US-10-040-862-9051 | Sequence 9051, App |
| 25 | 206 | 19.0 | 390 | 15 | US-10-057-475B-9051 | Sequence 9051, App |
| 26 | 206 | 19.0 | 390 | 15 | US-10-154-884B-9051 | Sequence 189, App |
| 27 | 198 | 18.2 | 821 | 15 | US-10-388-934-189 | Sequence 189, App |
| 28 | 192 | 17.7 | 674 | 14 | US-10-257-826A-150 | Sequence 150, App |
| 29 | 192 | 17.7 | 556 | 9 | US-09-796-692-5061 | Sequence 5061, App |
| 30 | 175 | 16.1 | 556 | 14 | US-10-040-862-5061 | Sequence 5061, App |
| 31 | 175 | 16.1 | 556 | 15 | US-10-057-475B-5061 | Sequence 5061, App |
| 32 | 175 | 16.1 | 556 | 15 | US-10-154-884B-5061 | Sequence 5061, App |
| 33 | 167 | 15.4 | 445 | 15 | US-10-378-029-25 | Sequence 25, Appl |
| 34 | 166.5 | 15.3 | 476 | 10 | US-09-918-995-3385 | Sequence 3385, A |
| 35 | 164 | 15.1 | 553 | 9 | US-09-796-692-4504 | Sequence 4504, App |
| 36 | 164 | 15.1 | 553 | 14 | US-10-040-862-4504 | Sequence 4504, App |
| 37 | 164 | 15.1 | 553 | 15 | US-10-057-475B-4504 | Sequence 4504, App |
| 38 | 164 | 15.1 | 553 | 15 | US-10-154-884B-4504 | Sequence 4504, App |
| 39 | 162.5 | 15.0 | 572 | 9 | US-09-864-761-9416 | Sequence 9416, App |
| 40 | 157.5 | 14.5 | 328 | 9 | US-09-796-692-5737 | Sequence 5737, App |
| 41 | 157.5 | 14.5 | 328 | 14 | US-10-040-862-5737 | Sequence 5737, App |
| 42 | 157.5 | 14.5 | 328 | 15 | US-10-057-475B-5737 | Sequence 5737, App |
| 43 | 157.5 | 14.5 | 328 | 15 | US-10-154-884B-5737 | Sequence 5737, App |
| 44 | 154 | 14.2 | 456 | 14 | US-10-029-386-25187 | Sequence 25187, A |
| 45 | 122 | 11.2 | 505 | 9 | US-09-796-692-4443 | Sequence 4443, App |

ALIGNMENTS

RESULT 1
US-09-966-880A-7
Sequence 7, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Homjo, Tasuku
APPLICANT: Muramatsu, Masamichi
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966, 880A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 2818
TYPE: DNA

ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (80)...(673)
 FEATURE:
 NAME/KEY: 5'UTR
 LOCATION: (1)...(79)
 FEATURE:
 NAME/KEY: 3'UTR
 LOCATION: (677)...(2818)
 US-09-966-880A-7

Alignment Scores:

| Pred. No.: | Length: | Matches: | Conservative: |
|------------------------|---------|-------------|---------------|
| 2.93e-140 | 2818 | 1086.00 | 198 |
| Percent Similarity: | 100.00% | Mismatches: | 0 |
| Best Local Similarity: | 100.00% | Indels: | 0 |
| Query Match: | 9 | Gaps: | 0 |

US-09-966-880A-8 (1-198) x US-09-966-880A-7 (1-2818)

QY 1 MetaspSerLeuMetaspSerLeuTyrglnPheLeuValArgTTP 20
 DB ATGACAGCCTCTGATGACCGGAGAGTTCTTACCAATTCAAAATGTCCTGG 139
 QY 21 AlalySGLYARGAGLUTHTYrLeuCYSTYrValValLysARGASerAlaThr 40
 DB GCTAAGGATCGCGTGAACCTGCTGCTAGCTAGAGAGAGCGTGCAGTCA 199
 QY 41 SerPheSerLeuAspPheGlyTYrLeuARGANLYSANGLYCYSHISVALGULEU 60
 DB TCCTTTTCACTGAGACTTGTGTTATCTTGCATAAGAGCGCTGCCAGTGTCTC 259
 QY 61 PheLeuArgTYrLISerAspTTPAspLeuAspProGlyARGCYSTYrValThTTP 80
 DB TTCCTCCGCTACATCTCGGACTGCGAGCTGAGCCTGCGCTGCTGCTGCTG 319
 QY 81 PheThrSerTTPSerProCYSTYrAspCYSAIARGHISVALAAspPheLeuArgGly 100
 DB TTCCACTCTCTGAGCGCCCTGCTACAGCTGTGCCCAATGTGGCGACTTCTCGAGGG 379
 QY 101 AspProAsnLeuSerLeuArgTLePheThrAlaARGLeuTYrPheCYSGIAspArgLYS 120
 DB AACCCCAACCTCAAGTGTGAGATCTTCAACCGCGCTCTACTTCTGTGAGAGCCGAG 439
 QY 121 AlaGluProGluGlyLeuArgArgLeuHISARGAlaGlyValGlnIleAlaIleMetThr 140
 DB GGTGAGCCGAGAGGGGTGGCGGCTGACCGCGCGGGGTGCAAAATGACATCATGACC 499
 QY 141 PheLYAspTYrPheTYrCYSTTPAsnThrPheValGluAsnHISGluArgThPheLYS 160
 DB TTCAAAGATTATTTTACTGCTGGAATCTTTTGTGAAGAAACCAAGAAAGACTTTCAA 559
 QY 161 AlaTTPGluGlyLeuHISGluAsnSerValArgLysSerArgGluLeuArgArgIleLeu 180
 DB GCCTGGAGAGGGGTGCATAAATTCAGTTCCTTCACAGCACTTCGCGCATCTT 619
 QY 181 LeuProLeuTYrGluValAspAspLeuArgAspAlaPheArgThrLeuGlyLeu 198
 DB TTGCCCTGTATGAGGTGATGACTTACAGAGAGCATTTGTACTTTGGGACTT 673

RESULT 2

US-09-966-880A-1
 Sequence 1, Application US/09966880A
 Patent No. US20020164743A1

GENERAL INFORMATION:
 APPLICANT: Honjo, Tasuku
 APPLICANT: Muramatsu, Masamichi
 TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
 FILE REFERENCE: 06501-088001
 CURRENT APPLICATION NUMBER: US/09/966, 880A
 CURRENT FILING DATE: 2001-09-28

PRIOR APPLICATION NUMBER: PCT/JP00/01918
 PRIOR FILING DATE: 2000-03-28
 PRIOR APPLICATION NUMBER: JP 11-371382
 PRIOR FILING DATE: 1999-12-27
 PRIOR APPLICATION NUMBER: JP 11-178999
 PRIOR FILING DATE: 1999-06-24
 PRIOR APPLICATION NUMBER: JP 11-87192
 PRIOR FILING DATE: 1999-03-29
 NUMBER OF SEQ ID NOS: 36
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1
 LENGTH: 2440
 TYPE: DNA
 ORGANISM: Mus musculus
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (93)...(686)
 FEATURE:
 NAME/KEY: 5'UTR
 LOCATION: (1)...(92)
 FEATURE:
 NAME/KEY: 3'UTR
 LOCATION: (690)...(2440)
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)...(2440)
 OTHER INFORMATION: n = A,T,C or G
 US-09-966-880A-1

Alignment Scores:

| Pred. No.: | Length: | Matches: | Conservative: |
|------------------------|---------|-------------|---------------|
| 1.76e-129 | 2440 | 1008.00 | 183 |
| Percent Similarity: | 95.94% | Mismatches: | 8 |
| Best Local Similarity: | 92.82% | Indels: | 0 |
| Query Match: | 9 | Gaps: | 0 |

US-09-966-880A-8 (1-198) x US-09-966-880A-1 (1-2440)

QY 1 MetaspSerLeuMetaspSerLeuTyrglnPheLeuValArgTTP 20
 DB ATGACAGCCTCTGATGACCGGAGAGTTCTTACCAATTCAAAATGTCCTGG 152
 QY 21 AlalySGLYARGAGLUTHTYrLeuCYSTYrValValLysARGASerAlaThr 40
 DB GCCAAGGACGGATGACAGCTACTGCTGCTAGAGAGAGAGATGTGCCACC 212
 QY 41 SerPheSerLeuAspPheGlyTYrLeuARGANLYSANGLYCYSHISVALGULEU 60
 DB TTCCTGCTACTGAGCTTGGCGACCTTGCACCAAGTGTGGCTGCCACTGGAATTGTG 272
 QY 61 PheLeuArgTYrLISerAspTTPAspLeuAspProGlyARGCYSTYrValThTTP 80
 DB TTCTTAGCTTCAATCTCAAGCTGAGACTGAGCCGGCGGTGTTACCGGCTCACTGG 332
 QY 81 PheThrSerTTPSerProCYSTYrAspCYSAIARGHISVALAAspPheLeuArgGly 100
 DB TTCACTCTCTGAGCGCCGCTGCTACAGCTGTGCCCAAGTGTGGCTGCAAGTGTG 392
 QY 101 AspProAsnLeuSerLeuArgTLePheThrAlaARGLeuTYrPheCYSGIAspArgLYS 120
 DB AACCTTAACTCTGAGAGATTTTACCGCGCGCTCTACTTCTGTGAAGACCGCAGG 452
 QY 121 AlaGluProGluGlyLeuArgArgLeuHISARGAlaGlyValGlnIleAlaIleMetThr 140
 DB GCTGAGCTTGAAGGGGTGGAGAGCTGACCGCGCTGGGGTCCAGATCCGATCATGACC 512
 QY 141 PheLYAspTYrPheTYrCYSTTPAsnThrPheValGluAsnHISGluArgThPheLYS 160
 DB TTCAAACATATTTTACTGCTGGAATCTTTTGTGAAGAAATGTGAAGAACTTTCAA 572
 QY 161 AlaTTPGluGlyLeuHISGluAsnSerValArgLysSerArgGluLeuArgArgIleLeu 180

D5 573 GCCTGGAGAGGCGCTACATGAAAAATTCTGTCCGGCTACACGACCACTTCGGGCGCATCCTT 633
QY 181 LeuProLeuTyrGluValAspAspLeuArGAspAlaPheArgThrLeuGly 197
D5 633 TTGGCCCTTGTACGAGTCATGACTTCGCGAGATGCACTTTCCGATTTCTTGGGA 683

RESULT 3

```

US-09-966-880A-10
; Sequence 10 Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masaamichi
; TITLE OF INVENTION: NOVEL CYCLINDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 6564
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-966-880A-10

```

| | |
|------------------------|---------|
| Alignment Scores: | |
| Pred. No.: | 3,9e-78 |
| Score: | 644 |
| Percent Similarity: | 24.02% |
| Best Local Similarity: | 24.02% |
| Query Match: | 59.35% |
| DB: | 9 |
| | |
| Length: | 656 |
| Matches: | 177 |
| Conservative: | 0 |
| Mismatches: | 2 |
| Indels: | 560 |
| Gaps: | 2 |

US-09-966-880A-8 (1-198) X US-09-966-880A-10 (1-6564)

| | | | |
|----|------|--|-------|
| QY | 3 | SerialNumMeAsnArgArglySphLeuYrGlnPheIysAsnValArgTrrAlaIys | 22 |
| Db | 1063 | AGCCCTTGATGAACCGAGGAAGTTCTTTACCAATTCAAAAAGTCGCGGGCTAAG | 11222 |
| QY | 23 | GIArgArgGluThrTyrLeuCysTyrValValIysArgArgSerpAlaThrSerPhe | 42 |
| Db | 1123 | GGTCGGCGGAACTACCTGCTGCTACGTAAGTGAAGAGCGCTGACAGTGCCTACATCCTTT | 1182 |
| QY | 43 | SerialSppHaeGlyTyrLeuArgAsnLys | 52 |
| Db | 1133 | TCACTGGACTTTGGTTATCTTGCAATAA-GGATCAATTAAGTCAGCTTGGCAAGAG | 1241 |
| QY | 52 | ----- | 52 |
| Db | 1242 | TTTATGTCACACTGTGAGTGTCTTTAAGGCACTGTCGATGATTACTTCATCCTT | 1301 |
| QY | 52 | ----- | 52 |
| Db | 1302 | TTTTGGCATTTGTGTCTCATCACATTCCTCAATTCCTTTTTTATTCTTTTCCATG | 1361 |
| QY | 52 | ----- | 52 |
| Db | 1362 | TCCATGACCCATATTAGACATGGCCAAATATGTGATTTAATTCCTCCCGATTAATGC | 1421 |
| QY | 52 | ----- | 52 |
| Db | 1422 | TGGGACCCCTAATACCACTCCTTCCTTCAGTCCAAAGCAACGTGCTCCCAAACTGTTTA | 1481 |
| QY | 52 | ----- | 52 |

| | | | |
|----|------|--|------|
| Db | 1482 | CCAGCTTCCGACGACATCTGAATTGCCTTTGAGATTAAATTAGCTAAAGACATTTTATA | 1541 |
| QY | 52 | ----- | 52 |
| Db | 1542 | TGGAGAAATATTATCAGCTTGCCAGCAAAATTTTAAATGTGMAAAACAAATTGTGTC | 1601 |
| QY | 52 | ----- | 52 |
| Db | 1602 | TTAAGCATTTTGGAAAATTAAAGAGAGAAATTTGGAAAAAATTAACGGTGTTCAATT | 1661 |
| QY | 52 | ----- | 52 |
| Db | 1662 | CTGTTTTCAAAATGATTTCTTTTCCCTCCTACTACATGGGTGTAGCCAGTGAATACA | 1721 |
| QY | 52 | ----- | 52 |
| Db | 1722 | TTCAACATGTGATCCCCAGAAAACTCAGAGAAAGCTCGGTGATGATTAATTAATGA | 1781 |
| QY | 52 | ----- | 52 |
| Db | 1782 | TCCTTCGGCTACCCGAGAGAAATTACATTTCCAGAGACTTCTTACCAAAAATCCAGATGG | 1841 |
| QY | 52 | ----- | 52 |
| Db | 1842 | GTTTACATAAACTTCTGCCATGGGTATCTCTCTCTAACAAGCTGTGAGCTCGGG | 1901 |
| QY | 52 | ----- | 52 |
| Db | 1902 | CTTGTTGGAATCTCAGGGAAGACCTCGTGGGTGGAAGTCATCGTCTGGCTCGTTGTTT | 1961 |
| QY | 52 | ----- | 52 |
| Db | 1962 | GATGTTAATATACCATGCAATTTCTTGCTTACATTTGATTAATGAATACATCCAACT | 2021 |
| QY | 52 | ----- | 52 |
| Db | 2022 | CCTTCCTATTCGGTGAATGACATTCATTTTCAGAAAGCTTGATTTATCAAGACT | 2081 |
| QY | 52 | ----- | 52 |
| Db | 2082 | TTGATTTACTTCTCATGGAGTGCCTATTACTTCTTACAAATCCATCTGTCTGTTT | 2141 |
| QY | 52 | ----- | 52 |
| Db | 2142 | ACCAAAATCTATTTCCCTTTTCAGATCTCCAAATGTCTCATAACTGTCTGCTT | 2201 |
| QY | 52 | ----- | 52 |
| Db | 2202 | CAACTGTGTGCCAGTATATTTCACAATGTTTACATCAACAGGCACTTACGCATTT | 2261 |
| QY | 52 | ----- | 52 |
| Db | 2262 | TCCTTTCAAAAGGTGCAAAAAGCACTTCATAAACAATAATTAATCTTGGTGAAGTA | 2321 |
| QY | 52 | ----- | 52 |
| Db | 2322 | GTGTGATGCTGTTCTTCCCAACTCAGGCACTTGCTTCTCTCATTTCCAAAAACCA | 2381 |
| QY | 52 | ----- | 52 |
| Db | 2382 | TAGCTTCTCTCACTCTGCAAGACTAGTGTGCCAAGGTTCACTTACTGTGTT | 2441 |
| QY | 52 | ----- | 52 |
| Db | 2442 | GCTCTTTGACAGATTTGCTTAGGCTCTGTGTAACAAGAAGATAGCTCAAGACATCC | 2501 |
| QY | 52 | ----- | 52 |
| Db | 2502 | CCAAAGATCATTGCAAGAGACAATGACTAAGGCTACAGAGCCGCAATTAAGTCAAGTA | 2561 |
| QY | 53 | ----- | 62 |
| Db | 2562 | ATTTTACGCTGCTCTGTCTCTCTCCAGAGAGGCTGCCAGTGAATGTCTCTCTCTC | 2621 |

QY 63 ArgTyrTlIseSerAspTTPaPLeuAapProGlyArgCySerYrArGValThrTPheThr 82
 Db 2622 CGGTACATCTCGAGCTGGAGACCTAGACCCGCGCTGCTACCGCTGCACCTGGTTCACC 2681
 QY 83 SerTTPSerProCySerYrAapCySAIArGHisValAlaAapPheLeuArgGlyAapPro 102
 Db 2682 TCTGGAGAGCCCTGCTACGACTGTGCGGACATGTGGCCACTTTCTGGAGGGAACCCC 2741
 QY 103 AsnLeuSerLeuArgIlePheThrAlaArgLeuTyrPheCysGluAapArgIlySAIAglu 122
 Db 2742 AACCTCAGCTCGAGGAGCTTCAACCGCGCGCTCTACTTGTGAGGAGCCGCAAGCTGAG 2801
 QY 123 ProGluGlyLeuArgIlePheThrAlaArgIlyValGlnIleAlaIleValThrPheLys 142
 Db 2802 CCCGAGGGGCTGCGCGGCTGCAACCGCGGGGTGCAATAGCCATCATGACCTTCAA- 2860
 QY 142 ----- 142
 Db 2861 AGGTGCGAAGGGCTTCCGCGAGGCGAGCTGCAAGCCCGCATTGGGATTGGAGT 2920
 QY 142 ----- 142
 Db 2921 CGGAATGATAGTTAGTGGGAAAGCTCGAGGGGAAGAGTGGCGGGGATTCTGTCTCA 2980
 QY 142 ----- 142
 Db 2981 CCTCGAGCGGAATTAAGATTAGAAGAGAGAAAGATGATGCTCAGACAGACAG 3040
 QY 142 ----- 142
 Db 3041 GCCCGAGAAATGAGAAATGGGCGCAGGGTGTCTTCCCTCGATTGGAACCTG 3100
 QY 143 ----- AspTyr 144
 Db 3101 AACTGCTCTTACCCCCATATCCCGCTTTTTCCTTTTTCCTTTTGAAGATTAT 3160
 QY 145 PheTyrCysTTPaPLeuThrPheValGluAapPheLysAlaTTPGluGly 164
 Db 3161 TTTTACTGTGGAATACCTTTGTAGAAACCGAAGAACCTTCAAGCTGGGAAAGG 3220
 QY 165 LeuHisGluAapSerValArgLeuSerArgGlnLeuArgIlyLeuLeu 181
 Db 3221 CCGCATGAAATTCAGTCTGTCTCCAGACAGCTTCGCGCATCTCTTTG 3271
 RESULT 4
 US-09-966-880A-35
 / Sequence 35, Application US/09966880A
 / Patent No. US20020164743A1
 / GENERAL INFORMATION:
 / APPLICANT: Honjo, Tasuku
 / APPLICANT: Mutamatsu, Masamichi
 / TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
 / FILE REFERENCE: 06501-088001
 / CURRENT APPLICATION NUMBER: US/09/966,880A
 / CURRENT FILING DATE: 2001-09-28
 / PRIOR APPLICATION NUMBER: PCT/JP00/01918
 / PRIOR FILING DATE: 2000-03-28
 / PRIOR APPLICATION NUMBER: JP 11-371382
 / PRIOR FILING DATE: 1999-12-27
 / PRIOR APPLICATION NUMBER: JP 11-178999
 / PRIOR FILING DATE: 1999-06-24
 / PRIOR APPLICATION NUMBER: JP 11-871192
 / PRIOR FILING DATE: 1999-03-29
 / NUMBER OF SEQ ID NOS: 36
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 35
 / LENGTH: 11204
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 US-09-966-880A-35

Pred. No.: 8,85e-78 Length: 11204
 Score: 644.50 Matches: 177
 Percent Similarity: 24.02% Conservative: 0
 Best Local Similarity: 24.02% Mismatches: 2
 Query Match: 59.35% Indels: 560
 DB: 9 Gaps: 2
 US-09-966-880A-8 (1-198) x US-09-966-880A-35 (1-11204)
 QY 3 SerLeuMetLeuAapArgIlyPheLeuTyrGlnPheYsAenValArgTTPAlaLys 22
 Db 6278 AGCCTCTGATGAACCGGAGAGATTCTTTACCAATTCAAAATGTCGCTGGGCTTAG 6337
 QY 23 GlyArgGluThrTyrLeuCySerYrValValIlyArgArgAapSerAlaThrSerPhe 42
 Db 6338 GGTGGGCTGAGACCTACCTGCTAGCTAGTGAAGAGCGGTGACAGTGTACATCTCTT 6397
 QY 43 SerLeuAapPheGlyTyrLeuArgAapLys ----- 52
 Db 6398 TCACTGACTTGGTTATCTTCGCAATPA-GATATCAATTAAGTCAGCTTTGCAAGCAG 6456
 QY 52 ----- 52
 Db 6457 TTAAATGCTCACTGTAGTCTTTAGAGCCACCTGCTAGTATTAATTCATCTT 6516
 QY 52 ----- 52
 Db 6517 TTTTGGCAATTGTGCTCTATTCATCAATTCCTAAATCTTTTATTTCTTTTCATG 6576
 QY 52 ----- 52
 Db 6577 TCCATGACCCATATTAGACATGCGCCAAATATGTATTAATCTCCCAATATGC 6636
 QY 52 ----- 52
 Db 6637 TGGGACCCCTAATACCATCTCTCTTCAAGTGCAGAACAACTGCTCCAACTGTTA 6696
 QY 52 ----- 52
 Db 6697 CCAGCTTCTCAGCATCTGAATTCCTTGAATTAATTAAGCTAAAGCATTTTATA 6756
 QY 52 ----- 52
 Db 6757 TGGGAATATTATACGCTTGTCCAGAAATTTAAATGGAACAAATTTGTCTC 6816
 QY 52 ----- 52
 Db 6817 TTAAGCATTTTGAATTAAGAGAGAAATTTGGAAAAATTAACGTTGTTCAAT 6876
 QY 52 ----- 52
 Db 6877 CTGTTTCCAAATGATTCTTTCTCTACTACATGAGTGTAGGCCATGATACA 6936
 QY 52 ----- 52
 Db 6937 TTCAACATGTGATCCCCAGAAACTCAGAGAAAGCTCGGTGATGATTAAATTA 6996
 QY 52 ----- 52
 Db 6997 TCTTGGGCTACCGGAGAGATTAATTCAGAGAGACTTTCACCAAAATCCAGATG 7056
 QY 52 ----- 52
 Db 7057 GTTACATAAATCTGCGCCAGTATCTCTCTCTCTCTCAACAGCTGACGTGG 7116
 QY 52 ----- 52
 Db 7117 CTGTGTGAATCTCAGGAAACATCCGTGGGTGAAAGGTGATGATGCTGCTGTGTT 7176
 QY 52 ----- 52
 Db 7177 GATGTTATATTACATGCAATTTCTTCTGCTACATTTGATTAATACATCCCATCT 7236

```
QY 52 ----- 52
Db 7237 CCTTCATTTGGTGACATGACATCTTATTTCAGAAAGGCTTGATTATTCAGACACT 7236
QY 52 ----- 52
Db 7237 TTCAATTACTTCTCATGGCAGAGTCTTACTTCTCTTACAAATACCATCTGTCTGCTT 7236
QY 52 ----- 52
Db 7357 ACCAAATCTATTTCCTTTTCAGATCTCCCAATGTCCTCATTAATCTGCTGCT 7416
QY 52 ----- 52
Db 7417 CCACCTAGTGTCCAGGATATTTTCACAAATGTTACATCAAGGACCTTACCATTT 7476
QY 52 ----- 52
Db 7477 TCCTTCTCAAAAGTGCAAAAAGCACTTCATAAACAATTAATCTTCGTAGAGTA 7536
QY 52 ----- 52
Db 7537 GTGTGATGCTGCTTCTCCCACTCAGGCACTTCTCTCTCATTCACAAAACCA 7596
QY 52 ----- 52
Db 7597 TAGCTTCTTCACTCTGACAGACTAGTCTGCCAAGGTTCACTCTACTCTGTGT 7656
QY 52 ----- 52
Db 7657 GCTCTTTGAGCAAGTCTTACCTCTCTGTAAACAAGGACATAGCTGCAAGCATCC 7716
QY 52 ----- 52
Db 7717 CCAAAAGATCATTCGAGGACATGACTAGGCTACCAAGCCGCAATTAAGCTAGTA 7776
QY 7777 ATTTAGCGTGGTCTCTCTCTCTCTCTCCAGAACGGCTGCGACGTGGAATGCTCTCTC 7836
QY 63 ArgTYRILESerASPTrpASPLeuASPProGlyARGCysTYRARGValIThrTrpPheThr 82
Db 7837 CGTACATCTCGGACTGTGAGCTTACCCCTGCGCTGCTACCTCTGTGAGACCGCAAGCTGAG 7896
QY 83 SerTrpSerProCysTYRASPValIArgHisValIAlaASPLeuArgGlyASPPro 102
Db 7897 TCTTGAGACCCCTGCTACGACTGTGCCGACATGTGGCCGACTTCTTCCAGGGAACCCC 7956
QY 103 AsnLeuSerLeuArgIlePheThrAlaArgLeuTYRPhCysGluASPArgIleGlu 122
Db 7957 AACCTCACTCGAGAGATCTTACCGCGCGCTCTACTCTGTGAGACCGCAAGCTGAG 8016
QY 123 ProGluGlyLeuArgArgLeuHisArgAlaGlyValGlnIleAlaIleMetThrPheLeys 142
Db 8017 CCGGAGGGGCTGCGCGGCTGTGACCGCGCGGTGCATAATGACCATCATGACCTTCAA- 8075
QY 142 ----- 142
Db 8076 AGGTGCGAAAGGCTTCCGGCGAGGCGGCACTGACGACGCCGCAATTCGGGATTGGCATG 8135
QY 142 ----- 142
Db 8136 CGGAATGAATGATTAGTGGGGAAGCTGAGGGGAAGAAGTGGGGGAGATTCTGTTCA 8195
QY 142 ----- 142
Db 8196 CCTTGAAGCCGAATTAAGATTGAAGACAGAAAGAGTGAATGGCTCAGAGACAG 8255
QY 142 ----- 142
Db 8256 GCCCGAGGAAATGAGAAATGAGGCGCAGGGTGTCTCTTCCCTCGATTGGAACCTG 8315
QY 143 ----- 144
AspTYR 144
```

```
Db 8316 AACTGCTTCTACCCCATATCCCGCCTTTTTCCTTTTCTTTTCTTTTGAAGATTAT 8375
QY 145 PheTYRQysTrpAsnThrPheValIgluAsnHisGluArgIlePheLeysAlaTrpGluGly 164
Db 8376 TTTTACTGCTGGAATTAATCTTTGTAGAAAACCAAGAAAGACTTTCAAAAGCCCTGGAGAGG 8435
QY 165 leuHisGluAsnSerValArgLeuSerArgGluLeuArgArgIleLeuLeu 181
Db 8436 CTGCATGAATAATTCAGTTGTCTCTCCAGACAGCTTCGGCGCATCTTTTG 8486

RESULT 5
US-09-966-880A-13
; Sequence 13, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 271
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-966-880A-13

Alignment Scores:
Pred. No.: 8,366-61 Length: 271
Score: 505.00 Matches: 90
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 46,50% Indels: 0
DB: 9 Gaps: 0

US-09-966-880A-8 (1-198) x US-09-966-880A-13 (1-271)
QY 53 AsnGlyCysHisValIgluLeuLeuPheLeuArgTYRILESerASPTrpASPLeuASPPro 72
Db 1 AACGGGTGCGCAGTGGAAATGCTCTTCCCGCTACATCTCGAGCTGGGACCTACACCT 60
QY 73 GlyARGCysTYRARGValIThrTrpPheThrSerTrpSerProCysTYRASPValIArg 92
Db 61 GGGCGGTGTACCGCGCTCACCTGTTCACCTCTGAGGCTCTGTACGACTGTGCCGA 120
QY 93 HisValIAlaASPLeuArgGlyAsnProAsnLeuSerLeuArgIlePheThrAlaArg 112
Db 121 CATGTGGCCGACTTTTCTGTGAGGGAACCCCAACCTCACTGATGAGATCTTACCGCGGCC 180
QY 113 LeuTYRPhCysGluAspArgIleValIArgIleProGluGlyLeuArgArgLeuHisArgAla 132
Db 181 CTCTACTTCTGTGAGGACCGCAAGGCTGAGCCGAGGGGCTGCGCGGCTGCAACCGCGCC 240
QY 133 GlyValGlnIleAlaIleMetThrPheLeys 142
Db 241 GGGGTGCAATAATGACCATCATGACCTTCAA 270

RESULT 6
US-10-460-923-1
; Sequence 1, Application US/10460923
; Publication No. US20040009951A1
; GENERAL INFORMATION:
```


DB 1053 ATTCAATTAATGACATACAGTAATTAAGCACTGTGGACACCTTTGTGACCAACG 1112
QY 156 GUAUgthrhphelysalaITpIGluGlyLeuHISgluanserValaGluLeuSerArgGln 175
DB 1113 GGATGTCCCTCCACGCCCTGGAGTGAAGTATGATGACAGACCCAAAGACTGAGTGGAGG 1172
QY 176 LeuArgArgIleLeu 180
DB 1173 CTCGGGGCCATTCTC 1187
RESULT 8
US-09-925-300-699
Sequence 699, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
APPLICANT: Steve Ruben
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: P101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 699
LENGTH: 987
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-300-699
Alignment Scores:
Pred. No.: 4,756-41 Length: 987
Score: 369.50 Matches: 79
Percent Similarity: 58.52% Conservative: 24
Best Local Similarity: 44.89% Mismatches: 64
Query Match: 34.02% Indels: 9
Gaps: 4
US-09-966-880a-8 (1-198) x US-09-925-300-699 (1-987)

QY 11 PheLeuTYrGlnPhelysAsnValaArgTrrpAlaLysGlyArgaGlnThrTYrLeuGys 30
DB 147 TTCTACTTCACATTTAAACCTATGGAGCAACAGATCGAAGCAAACTGCTGTGC 206
QY 31 TyrValValLys-----ArgArgAspSerAlaThrSerPheSerLeuAspPheGlyTyr 48
DB 207 TTTCACCGTGAAGATTAAGCGCCGCTCACTTCTCTCTGAGAGAG-----GGCGTC 260
QY 49 LeuArgAsn-----LysAsnGlyCysHisValaGluLeuLeuPheLeuArgTyr 64
DB 261 TTCCGAAACCAAGTGGATTCGAGCCCATTTGTATGACAGAAAGTCTCTCTTGG 320
QY 65 IleSerAspTrrpAspLeuAspProGlyArgCysTYrArgValThrTrrpPheThrSerTrrp 84
DB 321 TTCTGCGACGACATrCTCTCTCTTACCAAGTAAGTACCGGTCACTGTGACATCTTGG 380
QY 85 SerProCysTYrAspCysAlaArgHisValaAlaAspPheLeuArgGlyAsnProAsnLeu 104
DB 381 AGCCCTTCCAGACTGTGAGGGAGGTGGCCGACTTCTGCGCAGGACAGCAAGCTG 440
QY 105 SerLeuArgIlePheThAlaArgLeuTYrPheCysGluAspArgGlySAlaGluProGln 124
DB 441 AATTCACCATCTTCAACGCCCGCTTCACTACTTCAG-----TATCATGTTTACAGAG 497
QY 125 GlyLeuArgaGlyLeuHisArgAlaGlyValGlnIleAlaIleMetThrPheLeuAspTyr 144
DB 498 GGGCTCCGACAGCTGAGTCAAGAAAGGGGTCTGTGAGATCAAGACATTAAGAAATTT 557
QY 145 PheTYrCysTrrpAsnThrPheValGluAsnHisgluArgThrPheLysAlaITpIGluGly 164

DB 558 AATATTGTGTGGAAACCTTTGTGTACATGATATAGCAACCATTCACGCTTGAAGGA 617
QY 165 LeuHISgluanserValaArgLeuSerArgGlnLeuArgArgIleLeu 180
DB 618 TTTAAACCACTTTCGACTTCTTGAAAGAGCTACGGAGAGTCTC 665
RESULT 9
US-10-098-841-268
Sequence 268, Application US/10098841
Publication No. US20020197679A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenhua
APPLICANT: Asundi, Vinod
APPLICANT: Xu, Chongjun
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feiyang
APPLICANT: Chen, Rui-Hong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhimei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Qian, Xiaohong B.
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20020197679A1e1 Nucleic Acids and
FILE REFERENCE: 784CIP2
CURRENT APPLICATION NUMBER: US/10/098,841
CURRENT FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 331
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 268
LENGTH: 1143
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (128)..(700)
US-10-098-841-268
Alignment Scores:
Pred. No.: 5,946-41 Length: 1143
Score: 369.50 Matches: 79
Percent Similarity: 58.52% Conservative: 24
Best Local Similarity: 44.89% Mismatches: 64
Query Match: 34.02% Indels: 9
Gaps: 4
US-09-966-880a-8 (1-198) x US-10-098-841-268 (1-1143)

QY 11 PheLeuTYrGlnPhelysAsnValaArgTrrpAlaLysGlyArgaGlnThrTYrLeuGys 30
DB 176 TTCTACTTCACATTTAAACCTATGGAGCAACAGATCGAAGCAAACTGCTGTGC 235
QY 31 TyrValValLys-----ArgArgAspSerAlaThrSerPheSerLeuAspPheGlyTyr 48
DB 236 TTTCACCGTGAAGATTAAGCGCCGCTCACTTCTCTCTGAGAGAG-----GGCGTC 289
QY 49 LeuArgAsn-----LysAsnGlyCysHisValaGluLeuLeuPheLeuArgTyr 64
DB 290 TTCCGAAACCAAGTGGATTCGAGCCCATTTGTATGACAGAAAGTCTCTCTTGG 349
QY 65 IleSerAspTrrpAspLeuAspProGlyArgCysTYrArgValThrTrrpPheThrSerTrrp 84

Db 350 TTCTGGACGACATCTGCTCTTAACACAAAGTACAGGTCCAGCTGTACATCTTGG 409
 QY 85 SerProCyseTyAspCyAlaArgHisValAlaAspPheLeuArgGlyAsnProAsnLeu 104
 Db 410 AGCCCTTGGCCAGACCTGTGCGAGGGAGGTGCGGAGCTTCTGCGCAGACAGCAAGTGTG 469
 QY 105 SerLeuArgIlePheThrAlaArgLeuTyPheCyGluAspArgGlyAlaGluProGlu 124
 Db 470 ATCTACACATCTTCAACCCCGCTCTACTACTTCCAG---TATCCATGTACACAGAG 526
 QY 125 GlyLeuArgArgLeuHisArgAlaGlyValGlnIleAlaIleMetThrPheLeuAspTy 144
 Db 527 GGGCTCCGACGCTGTGCTGCGAGGGGTGCTGTGAGACATGACTATGAAGATTTT 586
 QY 145 PheTyCyseTyAsnThrPheValGluAsnHisGluArgThrPheValAlaTrpGluGly 164
 Db 587 AAATATTGTGGGAAAACCTTGTGTACATGATGATGAGCCATTGACCTTGAAGAGGA 646
 QY 165 LeuHisGluAsnSerValArgLeuSerArgGlnLeuArgArgIleLeu 180
 Db 647 TTAATAACCACTTTCGACTTCTGAAGAGGCTACGCGAGAGTCTC 694
 RESULT 10
 US-09-880-107-3282
 / Sequence 3282, Application US/09880107
 / Patent No. US20020142981A1
 / GENERAL INFORMATION:
 / APPLICANT: Horne, Darci T.
 / APPLICANT: Vockley, Joseph G.
 / APPLICANT: Scherf, Uwe
 / APPLICANT: Gene Logic, Inc.
 / TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
 / FILE REFERENCE: 44921-5028-WO
 / CURRENT APPLICATION NUMBER: US/09/880,107
 / PRIOR FILING DATE: 2001-06-14
 / PRIOR APPLICATION NUMBER: US 60/211,379
 / PRIOR FILING DATE: 2000-06-14
 / PRIOR APPLICATION NUMBER: US 60/237,054
 / PRIOR FILING DATE: 2000-10-02
 / NUMBER OF SEQ ID NOS: 3950
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 3282
 / LENGTH: 1348
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURE:
 / OTHER INFORMATION: Genbank Accession No. US20020142981A1 U03891
 / NAME/KEY: unsure
 / LOCATION: (1)-(1348)
 / OTHER INFORMATION: n = a o r c o r g o r t
 US-09-880-107-3282
 Alignment Scores:
 Pred. No.: 5,25e-40 Length: 1348
 Score: 363.50 Matches: 83
 Percent Similarity: 58.12% Conservative: 28
 Best Local Similarity: 43.46% Mismatches: 57
 Query Match: 33.47% Indels: 23
 Gaps: 7
 US-09-966-880a-8 (1-198) x US-09-880-107-3282 (1-1348)
 QY 5 LeuMetAsnArgArgIleGlyPheLeuTyGlnPheLeuAsnValArgTrpAlaLeuGlyArg 24
 Db 109 TTGATGATCCACACATATTCCTCACTTCACTTAAACAAT-----GGCATTGGAAG 159
 QY 25 ArgGlnThrTyLeuCyseTyValValIleArgArgAspSerAlaThrSerPheSerLeu 44
 Db 160 CATAGACCTTACTGTGTGCTGCGAGGTGAGCGCTGACATGAGCACTCGGTCAAGATG 219
 QY 45 Asp-----PheGlyTyTrpLeuArgAsn-----LysAsn-----GlyCys 54
 Db 220 GACACAGACAGGGGCTTTTCAACACACAGCGCTAAGATCTTCTGTGTGCTTTTACGGC 279

QY 55 CysHisValGluLeuLeuPheLeuArgTyrlleSerAspTrpAspLeuAspProGlyArg 74
 Db 280 CGCCATGGAGAGCTGGCGCTTCTTGACCTGGTCTTCTTGGAGTGGACCCGGCCAG 339
 QY 75 CysTyArgValThrTrpPheThrSerTrpSerProCysTyArg-----CysAlaArg 92
 Db 340 ACTTCACAGAGGTCATCTGTTCATCTCTGAGCCCTGCTTCTCTGGGGCTGTGCCGG 399
 QY 93 HisValAlaAspPheLeuArgGlyAsnProAsnLeuSerLeuArgIlePheThrAlaArg 112
 Db 400 GAAGTCCGTGCTTCTTCAAGAAACACACAGTGAAGTGGCATCTTCTGCCCCC 459
 QY 113 LeuTyPheCyGluAspArgGlyAlaGluPro-----GluGlyLeuArgArgLeu 129
 Db 460 ATCTAT-----GATTACGACCCCTATATAGAGGAGCGCTGCAAAATGCTG 504
 QY 130 HisArgAlaGlyValGlnIleAlaIleMetThrPheLeuAspTyTrpTyTrpAsn 149
 Db 505 CGGATGCTGGGGCCCAAGTCTCCATCATGATGATGATTAAGCATCTCTGGAC 564
 QY 150 ThrPheValGluAsnHisGluArgThrPheValAlaTrpGluGlyLeuHisGluAsnSer 169
 Db 565 ACCTTTGGACACACAGAGGATGTCCTTCCAGCCCTGGAGTGAATGATGACACAGC 624
 QY 170 ValArgLeuSerArgGlnLeuArgArgIleLeu 180
 Db 625 CAAGCCCTGAGTGGAGGCTGCGGCGCATTTCTC 657
 RESULT 11
 US-10-247-671-62
 / Sequence 62, Application US/10247671
 / Publication No. US20030194721A1
 / GENERAL INFORMATION:
 / APPLICANT: Mikita, Thomas
 / APPLICANT: Shiftman, Dov
 / APPLICANT: Porter, Gordon, J.
 / APPLICANT: Kaser, Matthew R.
 / TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
 / FILE REFERENCE: PA-0050 US
 / CURRENT APPLICATION NUMBER: US/10/247,671
 / PRIOR FILING DATE: 2002-09-18
 / PRIOR APPLICATION NUMBER: 60/323,784
 / PRIOR FILING DATE: 2001-09-19
 / NUMBER OF SEQ ID NOS: 186
 / SOFTWARE: PERL Program
 / SEQ ID NO 62
 / LENGTH: 823
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURE:
 / NAME/KEY: misc.feature
 / OTHER INFORMATION: Incyte ID No. US20030194721A1 135626CB1
 US-10-247-671-62
 Alignment Scores:
 Pred. No.: 3,58e-38 Length: 823
 Score: 348.00 Matches: 77
 Percent Similarity: 59.41% Conservative: 24
 Best Local Similarity: 45.29% Mismatches: 49
 Query Match: 32.04% Indels: 20
 Gaps: 6
 US-09-966-880a-8 (1-198) x US-10-247-671-62 (1-823)
 QY 26 GluTrpTyLeuCyseTyValValIleArgArgAspSerAlaThrSerPheSerLeuAsp 45
 Db 3 AAGACTTACTTGTGCTGAGAGTGAAGCGCTGACATGACACTCGGTCAAGATGAC 62
 QY 46 -----PheGlyTyTrpLeuArgAsn-----LysAsn-----GlyCys 55
 Db 63 CAGCACAGGGGCTTTTCAACACACAGCGCTAAGATCTTCTGTGTGCTTTTACGGCGC 122

QY 56 HisValGlnIleuIeuPheIeuIuArgTyrLiesAspTrpAspLeuAspProGluValGlyCys 75
Db 123 CATTGGAGAGCTGGCTTCTTGAGACCTGGTCTCTTGACAGTGGACCGGGCCCAATC 182
QY 76 TyrArgValThrTrpPheThrSerTrpSerProCysTyrAsp-----CysAlaArgHis 93
Db 183 TACAGGGTCACTTGTTTCATCTCTTGGACCCCTGCTTCTCGGGGCTGGCGGGGAA 242
QY 94 ValAlaAspPheIeuIuArgLysProAsnLeuSerLeuArgLLeuPheThrAlaArgLeu 113
Db 243 GTGGTGTCGTTCTCTTGAGAGAACACAGACGTGAGACGTGGCATTTGGCTGGCCGATC 302
QY 114 TyrPheCysIeuIuAspArgLysAlaGluPro-----GlnGluYLeuArgLeuHis 130
Db 303 TAT-----GATTATGACCCCTCATATAGAGAGCGCTCCAAATGCTGGC 347
QY 131 ArgAlaGluValGlnIleAlaIleAsnThrPheLysAspTyrPheTyrCysTrpAsnThr 156
Db 348 GATGCTGGGGGCCCAAGCTCCATCATAGACATGAGTGGAGTTCGTGGTGGACACC 407
QY 151 PheValGlnIuAsnHisGluArgThrPheLysAlaTrpGluGluYLeuHisGluAsnSerVal 170
Db 408 TTGTGTACCGCCAGGAGATGTCCCTTCCAGCCCTGGAGTGAAGTGAAGAACACAGCCAA 467
QY 171 ArgLysSerArgGlnIeuIuArgLLeuAla 180
Db 468 GCCCTGAGTGGAGGCTGGCGGCCCATTTCT 497

LENGTH: 148
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-966-880A-12

Alignment Scores:
 Pred. No.: 4.81e-27 Length: 148
 Score: 260.00 Matches: 49
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 23.94% Indels: 0
 DB: 9 Gaps: 0

US-09-966-880A-8 (1-198) x US-09-966-880A-12 (1-148)

| | | | |
|----|-----|--|-----|
| QY | 4 | LeuLeuMetAsnArgArgLysPheLeuTyrGlnPheLysAsnValArgTTPAlaLysGly | 23 |
| DB | 2 | CTCTTGATGAACCGAGGAGATTCTTTACCAATCAAAAATGCCGCTGGGCTAAGGGT | 61 |
| QY | 24 | ArgArgGluThrTyrLeuCysTyrValIleLysArgArgAspSerAlaThrSerPheSer | 43 |
| DB | 62 | CGGCGTGAAGACCTACTGTGCTACGTAGAGAGGCGTGAACAGTGTACATCCTTTCA | 121 |
| QY | 44 | LeuAspPheGlyTyrLeuArgAsnLys | 52 |
| DB | 122 | CTGGACTTTGGTTATCTTCGCATTAAG | 148 |

Search completed: March 13, 2004, 15:47:23
 Job time : 323 secs

ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,538
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051,571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3159:
SEQUENCE CHARACTERISTICS:
LENGTH: 795 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHEetical: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...795
SEQUENCE DESCRIPTION: SEQ ID NO: 3159:
US-09-107-532A-3159

Query Match
Best Local Similarity 11.2%; Score 33.6; DB 4; Length 795;
Matches 63; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 189 GCATATGGGCTACAGTACGAGAGAGAAATTAATCTTTATAGACCAAGTTATAGAG 248
DB 397 GCAACTCACTAAAGAAAGAAAGAACTGAAAGTTATGCGAAGCAAAAGATTGCAAG 456
QY 249 CACAAATTAAGTCTTTATTTGAAGATCTTACCTGCTGTTTCCAAATTCAG 300
DB 457 TAGCACTTAATCTCTATTAGATAGCATTTATCTTCTACAGATTTCAG 508

RESULT 3
US-09-780-172-18/c
Sequence 18, Application US/09780172
Patent No. 6607916
GENERAL INFORMATION:
APPLICANT: Robert McKay
APPLICANT: Susan M. Freier
TITLE OF INVENTION: ANTISENSE MODULATION OF CASEIN KINASE 2-ALPHA EXPRESSION
FILE REFERENCE: PUS-0159
CURRENT APPLICATION NUMBER: US/09/780,172
FILING DATE: 2001-02-08
NUMBER OF SEQ ID NOS: 96
SEQ ID NO 18
LENGTH: 63000
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-09-780-172-18

Query Match
Best Local Similarity 11.0%; Score 33; DB 4; Length 63000;
Matches 58.8%; Pred. No. 5.6;

Matches 57; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 136 ATAGTTCCATTACAGTACGATGATGATCTACTAGTGAATTAATAGATGATATG 195
DB 24931 ATAGTGCAGATGCAAGATGATGACATTAATTAAGATTAATAATAGATATATTC 24872
QY 196 GGTACAGTACGAGAGAGAAATTAATCTTTAATAGA 232
DB 24871 CAACCGAAGAGAGAGAAAGTAAATCATTAAGAA 24835

RESULT 4
US-09-616-289-49
Sequence 49, Application US/09616289
Patent No. 6632923
GENERAL INFORMATION:
APPLICANT: Lees, Robert S.
APPLICANT: Lees, Robert S.
APPLICANT: Law, Simon W.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/09/616,289
FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 09/517,849
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 08/979,608
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: US 60/031,930
PRIOR FILING DATE: 1996-11-27
PRIOR APPLICATION NUMBER: US 60/048,547
FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 49
LENGTH: 12619
TYPE: DNA
ORGANISM: Homo sapiens
US-09-616-289-49

Query Match
Best Local Similarity 10.8%; Score 32.4; DB 4; Length 12619;
Matches 90; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 106 TTGCTTAGATGATGATCCCTAGCAGCTGATGATGTTCCATTACAGTACGATGA 165
DB 9961 TTGTTACTTTTGATCCCGACATCTAGCAGATGATGATTAACAAGTATTCATA 10020
QY 166 TCTACTCAGTAAATTAATGATATGCGATATGGGCTACAGTACGAGAGAGAAATTAATCTT 225
DB 10021 AATATTTTGAACAGAAAGAACAGAGCTCAGATTCCTAATCTTGATCATTCCTC 10080
QY 226 TAATAGCAAGTTCTTAGAGAGACAAATTAAGTCTTTATTTAGAGATCTTAGCT 285
DB 10081 TATTTTCAATTAAGTACGATGATTAAGTTCTTAGTTCTTAGCTATTAACAATACCT 10140
QY 286 GTTTTC 291
DB 10141 TCTTTC 10146

RESULT 5
US-09-399-913-48
Sequence 48, Application US/09399913
Patent No. 6361971
GENERAL INFORMATION:
APPLICANT: Rhodes, Kenneth
APPLICANT: Betty, Maria
APPLICANT: Lang, Hui-Ping
APPLICANT: An, Wenglian
TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ. ID NO.: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-Fls
US-08-232-463-14

```

| | | | | |
|-----------------------|------------------|-----------------|-----------|--------------|
| Query Match | 10.7% | Score 32.2; | DB 1; | Length 7218; |
| Best Local Similarity | 3.8%; | Pred. No. 4.3; | | |
| Matches 10; | Conservative 14; | Mismatches 107; | Indels 0; | Gaps 0; |

[illegible]

RESULT 7
US-10-204-708-46/c
; Sequence 46, Application US/10204708
; Patent No. 6677731

Patent No. 6677723
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBERCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
TITLE OR INVENTION: by Assessing DNA Methylation
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30

US-08-960-048-3

Query Match 10.4%; Score 31.2; DB 3; Length 1063;
Best Local Similarity 57.0%; Pred. No. 4.2;
Matches 57; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 171 TCAGTGAATTAATAGATGATGCTACAGTAGAGAGAGAAATTAATCTTTAATA 230
Db 251 TAGTGAACAGATTAATAGCTATCATTTGATTAGAGTGAATCTAAATTCGAAAA 310
Qy 231 GACCAAGTTCTATGAGACACAAATTAAGTCTTTATT 270
Db 311 GTATTAATACTAAAGATTAAATGAAGAACATTAAAT 350

RESULT 10

US-09-838-586-3
; Sequence 3, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/02/US
; CURRENT APPLICATION NUMBER: US/09/838,586
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1063
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-09-838-586-3

Query Match 10.4%; Score 31.2; DB 4; Length 1063;
Best Local Similarity 57.0%; Pred. No. 4.2;
Matches 57; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 171 TCAGTGAATTAATAGATGATGCTACAGTAGAGAGAGAAATTAATCTTTAATA 230
Db 251 TAGTGAACAGATTAATAGCTATCATTTGATTAGAGTGAATCTAAATTCGAAAA 310
Qy 231 GACCAAGTTCTATGAGACACAAATTAAGTCTTTATT 270
Db 311 GTATTAATACTAAAGATTAAATGAAGAACATTAAAT 350

RESULT 11

US-09-751-389-3
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001067
; CURRENT APPLICATION NUMBER: US/09/751,389
; CURRENT FILING DATE: 2001-01-02
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: (1)...(786431)
; OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3

Query Match 10.3%; Score 31; DB 4; Length 786431;
Best Local Similarity 55.0%; Pred. No. 56;
Matches 61; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 114 CATTGATCGCTTACGACCTGCGATAGTTCCATTAACTAGGATGAAGTACTTACTCA 173
Db 697374 CTTTGAGCTCAATATAGATTTAGAGTCTTTTACAGTATGCTCAAACTAACAGCTGA 697433
Qy 174 GTGAATTAATTAATAGATGATGCTACAGTAGAGAGAGAAATTAATCT 224
Db 697434 GGGAAAAAGAGAAAAAGAAAAAGTGAAGAGAGAGATTAATCT 697484

RESULT 12

US-09-247-155-67/C
; Sequence 67, Application US/09247155A
; Patent No. 6312922
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET.021A
; CURRENT APPLICATION NUMBER: US/09/247,155A
; CURRENT FILING DATE: 1999-02-09
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-04-13
; EARLIER APPLICATION NUMBER: 60/096,116
; EARLIER FILING DATE: 1998-08-10
; EARLIER APPLICATION NUMBER: 60/099,273
; EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent. pm
; SEQ ID NO 67
; LENGTH: 854
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 186..443
; FEATURE:
; NAME/KEY: sig peptide
; LOCATION: 186..407
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: seq 3.9
; OTHER INFORMATION: seq ISCTCLLVTHC/IL
; FEATURE:
; NAME/KEY: polyA signal
; LOCATION: 827..832
; FEATURE:
; NAME/KEY: polyA site
; LOCATION: 839..854
US-09-247-155-67

Query Match 10.3%; Score 30.8; DB 4; Length 854;
Best Local Similarity 48.3%; Pred. No. 5.1;
Matches 86; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 42 AATCATTAAGTATAGATGATGCTATCAAGGCTTAGGAGAGAGAGACGACCTTAGC 101
Db 468 ATTCAACAAATTTGTCACAAATTCATGATTAAGATTAAGAGATTAAGCAATATCTTATA 409
Qy 102 TGATGCTTAGAGATGATGCTAGACCTGCGATAGTTCCATTAACTAGTAGGATGA 161
Db 408 TACATGTGTTAATTAATTAACAGATGATGATTAATTAAGAGGAGGAGATGA 349
Qy 162 AGTATCTACTAGTGAATTAATTAAGATGATGATGCTAGTAGAGAGAGAAATTA 219

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2004, 17:19:49 ; Search time 526.309 Seconds
(without alignments)
2098.537 Million cell updates/sec

Title: US-09-966-880A-9_COPY_1_300

Perfect score: 300
Sequence: 1 acagacgaatacatgctcca.....agcctgttcccaattccag 300

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US6_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US6_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being pinned, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|------------------------------|---------------------|
| 1 | 300 | 100.0 | 5514 9 US-09-966-880A-9 | Sequence 9, Appl |
| 2 | 277.2 | 92.4 | 461 15 US-10-027-632-297069 | Sequence 297069, Ap |
| 3 | 277.2 | 92.4 | 461 15 US-10-027-632-297070 | Sequence 297070, Ap |
| 4 | 277.2 | 92.4 | 461 15 US-10-027-632-297071 | Sequence 297071, Ap |
| 5 | 273 | 91.0 | 810 15 US-10-027-632-4416 | Sequence 4416, Ap |
| 6 | 273 | 91.0 | 810 15 US-10-027-632-4417 | Sequence 4417, Ap |
| 7 | 273 | 91.0 | 810 15 US-10-027-632-4418 | Sequence 4418, Ap |
| 8 | 40.2 | 13.4 | 1210 15 US-10-027-632-214410 | Sequence 214410, Ap |
| 9 | 39.6 | 13.2 | 737 15 US-10-027-632-8782 | Sequence 8782, Ap |
| 10 | 39.6 | 13.2 | 11216 10 US-09-764-872-490 | Sequence 490, Appl |
| 11 | 37.6 | 12.5 | 3673778 14 US-10-312-841-2 | Sequence 841, Ap |
| 12 | 37.4 | 12.5 | 600 14 US-10-029-386-5897 | Sequence 5897, Ap |
| 13 | 37.2 | 12.4 | 470 15 US-10-027-632-270562 | Sequence 270562, Ap |
| 14 | 37.2 | 12.4 | 8943 16 US-10-257-166-47 | Sequence 47, Appl |
| 15 | 36.2 | 12.1 | 637 15 US-10-027-632-247353 | Sequence 247353, Ap |

| | | | | |
|----|------|------|------------------------------|---------------------|
| 16 | 36 | 12.0 | 6242 14 US-10-311-455-2121 | Sequence 2121, Ap |
| 17 | 35.4 | 11.8 | 580 15 US-10-027-632-32647 | Sequence 32647, A |
| 18 | 35.2 | 11.7 | 525 15 US-10-027-632-190320 | Sequence 190320, A |
| 19 | 35.2 | 11.7 | 585 15 US-10-027-632-190321 | Sequence 190321, A |
| 20 | 35.2 | 11.7 | 1826 15 US-10-027-632-43119 | Sequence 43119, A |
| 21 | 35.2 | 11.7 | 1826 15 US-10-027-632-43120 | Sequence 43120, A |
| 22 | 35.2 | 11.7 | 3529 15 US-10-369-493-46354 | Sequence 46354, A |
| 23 | 35 | 11.7 | 366 9 US-09-960-352-1681 | Sequence 1681, Appl |
| 24 | 35 | 11.7 | 379 9 US-09-960-352-6621 | Sequence 6621, Ap |
| 25 | 35 | 11.7 | 389 9 US-09-960-352-14093 | Sequence 14093, A |
| 26 | 35 | 11.7 | 413 9 US-09-960-352-2490 | Sequence 2490, Ap |
| 27 | 35 | 11.7 | 414 9 US-09-960-352-2489 | Sequence 2489, Ap |
| 28 | 35 | 11.7 | 414 9 US-09-960-352-10772 | Sequence 10772, A |
| 29 | 35 | 11.7 | 416 9 US-09-960-352-8509 | Sequence 8509, Ap |
| 30 | 35 | 11.7 | 426 9 US-09-960-352-6031 | Sequence 6031, Ap |
| 31 | 34.8 | 11.6 | 607 14 US-10-196-846-5935 | Sequence 5935, Ap |
| 32 | 34.6 | 11.5 | 2000 9 US-09-938-842A-4728 | Sequence 4728, Ap |
| 33 | 34.6 | 11.5 | 2000 11 US-09-938-842A-4728 | Sequence 4728, Ap |
| 34 | 34.6 | 11.5 | 128779 14 US-10-081-337-38 | Sequence 38, Appl |
| 35 | 34.2 | 11.4 | 360 12 US-10-424-559-99764 | Sequence 99764, A |
| 36 | 34.2 | 11.4 | 382 9 US-09-960-352-3288 | Sequence 3288, Ap |
| 37 | 34.2 | 11.4 | 634 15 US-10-027-632-200698 | Sequence 200698, Ap |
| 38 | 34 | 11.3 | 403 9 US-09-960-352-275 | Sequence 275, Appl |
| 39 | 34 | 11.3 | 695 15 US-10-027-632-27031 | Sequence 27031, A |
| 40 | 34 | 11.3 | 1060 14 US-10-195-730-24 | Sequence 24, Appl |
| 41 | 34 | 11.3 | 3123 15 US-10-027-632-25969 | Sequence 25969, A |
| 42 | 34 | 11.3 | 3133 15 US-10-027-632-259970 | Sequence 259970, A |
| 43 | 34 | 11.3 | 3133 15 US-10-027-632-259971 | Sequence 259971, A |
| 44 | 34 | 11.3 | 3133 15 US-10-027-632-259972 | Sequence 259972, A |
| 45 | 34 | 11.3 | 3252 15 US-10-027-632-113786 | Sequence 113786, A |

ALIGNMENTS

RESULT 1
US-09-966-880A-9
Sequence 9, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Tetsu
APPLICANT: Muramatsu, Masamichi
TITLE OF INVENTION: NOVEL CITRIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966, 880A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178939
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 5514
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: intron
LOCATION: (1)...(1031)
FEATURE:
NAME/KEY: exon
LOCATION: (1032)...(1118)
FEATURE:
NAME/KEY: intron
LOCATION: (1119)...(5514)
US-09-966-880A-9
Query Match 100.0%; Score 300, DB 9, Length 5514;
Best Local Similarity 100.0%; Pred. No. 9, 1e-79;

| | Matches | 300; Conservative | 0; Mismatches | 0; Indels | 0; Gaps | 0; |
|----|---------|---|---------------|-----------|---------|----|
| QY | 1 | ACGAGCAATATACATGGTCCCAAGCTAGGGCTATTGATTTGAAATCATCAAGATATAGATG | 60 | | | |
| Db | 1 | ACGAGCAAAATACATGGTCCCAAGCTAGGGCTATTATATTGAAATCATCAAGATATAGATG | 60 | | | |
| QY | 61 | GTATCAAAAGGCTTTGAGGCGAGAGAGACACAGACCCCTAGCTCATGTCTTGACATTGCA | 120 | | | |
| Db | 61 | GTATCAAAAGGCTTTGAGGCGAGAGAGACACAGACCCCTAGCTCATGTCTTGACATTGCA | 120 | | | |
| QY | 121 | TCCTTAGCACCTTGAGCATAGTTTCCATTACAGTACGATGAAAGTATCTACTCAGTGAATA | 180 | | | |
| Db | 121 | TCCTTAGCACCTTGAGCATAGTTTCCATTACAGTACGATGAAAGTATCTACTCAGTGAATA | 180 | | | |
| QY | 181 | AATAGATGATATGAGGCTACAGTACGAGAGAGAAATPAAATCTTTATATAGACCAAGTTC | 240 | | | |
| Db | 181 | AATAGATGATATGAGGCTACAGTACGAGAGAGAAATPAAATCTTTATATAGACCAAGTTC | 240 | | | |
| QY | 241 | TATGAGAGCACAAATTAATAAGTCTTTATTTGAAAGATCTTAGCTGTTTTCCAAATTCAG | 300 | | | |
| Db | 241 | TATGAGAGCACAAATTAATAAGTCTTTATTTGAAAGATCTTAGCTGTTTTCCAAATTCAG | 300 | | | |

```

RESULT 2
US-10-027-632-297069
; Sequence 297069, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 297069
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-297069

Query Match      92.4%; Score 277.2; DB: 15; Length 461;
Best Local Similarity 99.3%; Pred. No. 1.7e-72;
Matches 276; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACAGAGCAATACATGCTCCAGCTAGGGCTATTGATTGTAATAATCATCAAGGTATGATG 60
Db      183 ACAGAGCAATACATGCTCCAGCTAGGGCTATTGATTGTAATAATCATCAAGGTATGATG 242

QY      61 GTATCAAAGGCTTGAGGCGAGAAAGAGACAGACACCTAGCTGATTGCTTAGCATTTGA 120
Db      243 GTATCAAAGGCTTGAGGCGAGAAAGAGACAGACACCTAGCTGATTGCTTAGCATTTGA 302

QY      121 TCCCTGACACCTGACATAGTTTCCATTAAACAGTGGGCGTAGAGATCTACTAGTGAATA 180
Db      303 TCCCTGACACCTGACATAGTTTCCATTAAACAGTGGGCGTAGAGATCTACTAGTGAATA 362

QY      181 AATGAGATGCAATATGGGCTACAGTAGAGAGAGAAATTAATCTTAATAGACCAATTC 240

```

Db 363 AATGATCATGTGGGCTACAGTGGAGAGAGAAATTTCTTTATAGACCAAGTTC 422
 Qy 241 TATGAGAGACAAATTTAAAGCTTTTATTTGAGATC 278
 Db 423 TATGAGAGACAAATTTAAAGCTTTTATTTGAGATC 460

```

RESULT 3
US-10-027-632-297070
? Sequence 297070, Application US/10027632
? Publication No. US20030204075A9
? GENERAL INFORMATION:
? APPLICANT: Wang, David G.
? TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
? FILE REFERENCE: 108827.129
? CURRENT APPLICATION NUMBER: US/10/027,632
? CURRENT FILING DATE: 2002-04-30
? PRIOR APPLICATION NUMBER: US 60/218,006
? PRIOR FILING DATE: 2000-07-12
? PRIOR APPLICATION NUMBER: US 60/199,676
? PRIOR FILING DATE: 2000-04-20
? PRIOR APPLICATION NUMBER: US 60/193,483
? PRIOR FILING DATE: 2000-03-29
? PRIOR APPLICATION NUMBER: US 60/185,218
? PRIOR FILING DATE: 2000-02-24
? PRIOR APPLICATION NUMBER: US 60/167,363
? PRIOR FILING DATE: 1999-11-23
? PRIOR APPLICATION NUMBER: US 60/156,358
? PRIOR FILING DATE: 1999-09-28
? PRIOR APPLICATION NUMBER: US 60/146,002
? PRIOR FILING DATE: 1999-08-09
? NUMBER OF SEQ ID NOS: 325720
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO: 297070
? LENGTH: 461
? TYPE: DNA
? ORGANISM: Human
US-10-027-632-297070

```

```

Query Match Similarity 92.4%; Score 277.2; DB 15; Length 461,
Best Local Similarity 99.3%; Pred. No. 1,7e-72;
Matches 276; Conservative 2; Mismatches 0; Indels 0; Gaps 0

QY 1 ACAGACGAATACATGTGTCACAGTGGGCTATTGATTTGAAAATTCATCAAGGTATGATG 60
DB 183 ACAGACGAATACATGTGTCACAGTGGGCTATTGATTTGAAAATTCATCAAGGTATGATG 242

QY 61 GTATCAAAAGCTTGAGGACAGAGAAGACAGACCCTAGCTCATGCTTGACATGCA 120
DB 243 GTATCAAAAGCTTGAGGACAGAGAAGACAGACCCTAGCTCATGCTTGACATGCA 302

QY 121 TCCCTTAGCACTGSCAATGTTCCATTACAGTAGACATGAAGTATCTACTCAGTGATTA 180
DB 303 TCCCTTAGCACTGSCAATGTTCCATTACAGTAGACATGAAGTATCTACTCAGTGATTA 362

QY 181 AATAGATCATATGAGGCTACAGTAGAGAGAGAAATTTATTAATAGACCAAGTTC 240
DB 363 AATAGATCATATGAGGCTACAGTAGAGAGAGAAATTTATTAATAGACCAAGTTC 422

QY 241 TATGAGAGACAAAATTTAAAGTCTTTTATTTGAAGATC 278
DB 423 TATGAGAGACAAAATTTAAAGTCTTTTATTTGAAGATC 460

RESULT 4
US-10-027-632-297071
; Sequence 297071, Application US /10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

```


QY 1 ACAGACCAATTCATGCTCCAGCTAGGGCTATTGATTTGAAATCATCAAGGTATGAG 60
Db 536 ACAGACCAATTCATGCTCCAGCTAGGGCTATTGATTTGAAATCATCAAGGTATGAG 595
QY 61 GTATCAAAAGGCTTAGGAGGAGAGAGAGACCTAGCTGATTCCTTAGCATTCGA 120
Db 596 GTATCAAAAGGCTTAGGAGGAGAGAGAGACCTAGCTGATTCCTTAGCATTCGA 655
QY 121 TCCCTAGACCTGGATGATTTCCATTACAGTAGAGATGATCTACTAGTGAATA 180
Db 656 TCCCTAGACCTGGATGATTTCCATTACAGTAGAGATGATCTACTAGTGAATA 715
QY 181 AATGAAATGCAATGAGGCTACAGTAGAGAGAGAAATTTAAATCTTTAATGACCAAGTTC 240
Db 716 AATGAAATGCAATGAGGCTACAGTAGAGAGAGAAATTTAAATCTTTAATGACCAAGTTC 775
QY 241 TATGAGACCAAAATTAAGTCTTTTATTGAG 275
Db 776 TATGAGACCAAAATTAAGTCTTTTATTGAG 810

RESULT 7

US-10-027-632-4418
; Sequence 4418, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4418
; LENGTH: 810
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-4418

Query Match 91.0%; Score 273; DB 15; Length 810;
Best Local Similarity 99.3%; Pred. No. 4e-71;
Matches 273; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACAGACCAATTCATGCTCCAGCTAGGGCTATTGATTTGAAATCATCAAGGTATGAG 60
Db 536 ACAGACCAATTCATGCTCCAGCTAGGGCTATTGATTTGAAATCATCAAGGTATGAG 595
QY 61 GTATCAAAAGGCTTAGGAGGAGAGAGACCTAGCTGATTCCTTAGCATTCGA 120
Db 596 GTATCAAAAGGCTTAGGAGGAGAGAGACCTAGCTGATTCCTTAGCATTCGA 655
QY 121 TCCCTAGACCTGGATGATTTCCATTACAGTAGAGATGATCTACTAGTGAATA 180
Db 656 TCCCTAGACCTGGATGATTTCCATTACAGTAGAGATGATCTACTAGTGAATA 715
QY 181 AATGAAATGCAATGAGGCTACAGTAGAGAGAGAAATTTAAATCTTTAATGACCAAGTTC 240
Db 716 AATGAAATGCAATGAGGCTACAGTAGAGAGAGAAATTTAAATCTTTAATGACCAAGTTC 775

QY 241 TATGAGACCAAAATTAAGTCTTTTATTGAG 275
Db 776 TATGAGACCAAAATTAAGTCTTTTATTGAG 810

RESULT 8

US-10-027-632-214410
; Sequence 214410, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 214410
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-214410

Query Match 13.4%; Score 40.2; DB 15; Length 1210;
Best Local Similarity 62.4%; Pred. No. 0.16;
Matches 63; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 15 GTTCCAAAGCTAGGGCTATTGATTTGAAATCATCAAGGTATGAGGCTG 74
Db 828 GTTCCAAAGCTAGGGCTATTGATTTGAAATCATCAAGGTATGAGGCTG 887
QY 75 AGGAGAGAGAGAGAGAGACCTAGCTGATTCCTTAGCA 115
Db 888 AGGAGAGAGAGAGAGAGACCTAGCTGATTCCTTAGCA 928

RESULT 9

US-10-027-632-8782/c
; Sequence 8782, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8782
LENGTH: 737
TYPE: DNA
ORGANISM: Human
US-10-027-632-8782

Query Match 13.2%; Score 39.6; DB 15; Length 737;
Best Local Similarity 66.3%; Pred. No. 0.19;
Matches 57; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 98 TAGCTGCACTTGGCTTACGATTCCTAGCAGCCTGGCATAGTTCCATTACAGTACGC 157
DB 385 TATCTATTGCTCTCATATGATCATCTCGATCGAAGACAGCTGTGCTTGCATTAGGC 326
QY 158 ATGAAGTATCTCATCGATGAATAAT 183
DB 325 ACTCAGTATTACTGATGAATAAT 300

RESULT 10
US-09-764-872-490
Sequence 490, Application US/09764872
Publication No. US20030050231A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA125
CURRENT APPLICATION NUMBER: US/09/764,872
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 957
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 490
LENGTH: 11216
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-872-490

Query Match 13.2%; Score 39.6; DB 10; Length 11216;
Best Local Similarity 66.3%; Pred. No. 0.75;
Matches 57; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 98 TAGCTGCACTTGGCTTACGATTCCTAGCAGCCTGGCATAGTTCCATTACAGTACGC 157
DB 10173 TATCTATTGCTCTCATATGATCATCTCGATCGAAGACAGCTGTGCTTGCATTAGGC 10332
QY 158 ATGAAGTATCTCATCGATGAATAAT 183
DB 10233 ACTCAGTATTACTGATGAATAAT 10258

RESULT 11
US-10-312-841-2/C
Sequence 2, Application US/10312841
Publication No. US20030186277A1
GENERAL INFORMATION:
APPLICANT: Epigenomics AG
TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
FILE REFERENCE: E01/1208/NO
CURRENT APPLICATION NUMBER: US/10/312,841
CURRENT FILING DATE: 2002-12-30
NUMBER OF SEQ ID NOS: 2
SEQ ID NO 2
LENGTH: 3673778
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
FEATURE:
NAME/KEY: unsure
LOCATION: (379615)
US-10-312-841-2

Query Match 12.5%; Score 37.6; DB 14; Length 3673778;
Best Local Similarity 61.0%; Pred. No. 55;
Matches 61; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 177 AATTAATGATTCATATGGGCTTACAGTACGAGAGAGATAATTAATCTTTAATAGCCAA 236
DB 50379 AAAAAACAATTCATATCTTTAATTCGAAAAATATATAATTTAATATACCA 50320
QY 237 GTTCTATGAGAGACAAAATTAAGTCTTTAATTTGAGA 276
DB 50319 TTAATAATACCATTAATAACAACTAATATACAAA 50280

RESULT 12
US-10-029-386-5897
Sequence 5897, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: ABOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 5897
LENGTH: 600
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
OTHER INFORMATION: MAP TO AF216808.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.98
OTHER INFORMATION: NT HIT: AL163267.2, EVALUE 1.50e-01
OTHER INFORMATION: EST_HUMAN HIT: BE243894.1, EVALUE 8.40e-01
US-10-029-386-5897

Query Match 12.5%; Score 37.4; DB 14; Length 600;
Best Local Similarity 54.8%; Pred. No. 0.77;
Matches 74; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 88 GCAAGACCTTACCTGATTCCTAGCATGATCCCTAGCAGCCTGGCATAGTTCCATT 147
DB 279 GCAGAAAGTAATATGCTTAGCATCTTGAATCCCTAGTACCTGAGACATGTCTAAT 338
QY 148 AACAGTGGCATGATGATCTACTCAGTGAATTAATGAATGATATGGCTTACAGTAG 207
DB 339 CTTAGTAGGATGATGATGCTATTATTAATAAATAATCATGATGATTAATAAGTAT 398
QY 208 AGAGAGAAATTAAT 222
DB 399 AAATGGAGATAAT 413

RESULT 13
US-10-027-632-270562/C
Sequence 270562, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129

```

RESULT 14
US-10-257-166-47/c
; Sequence 47, Application US/10257166
; Publication No. US2004002320A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Method and Nucleic Acids for Analysing the Methylation of
; TITLE OF INVENTION: Genes Implicated in Pharmacogenomics
; FILE REFERENCE: 5013.1011
; CURRENT APPLICATION NUMBER: US/10/257,166
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: PCT/EP01/07470
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-06-29
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 47
; LENGTH: 8943
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-257-166-47

```

```

Query Match Similarity 12.1%; Score 36.2; DB 15; Length 637;
Best Local Similarity 65.4%; Pred. No. 1.8; Matches 28; Indels 0; Gaps 0
Matches 53; Conservative 0; Mismatches 28; Indels 0; Gaps 0

QY 68 AGCCTTGAGGACGAGAAAGAGCAGAGACCTAGCTGCATGCTTATGACATGCAATCCCTG 127
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 469 AAAATGTATAGCTCCAGGAAGCAGAAACAGTCTCACTGCTTATAGCTGCATCCCTAG 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 128 CACCTGGCATAGTTTCATTA 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 409 CACCTGGCATGATGCTTGTGA 389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Search completed: March 12, 2004, 09:38:41
Job time : 533.309 secs

```


| | | | | | | |
|----|------|------|--------|------------------|--------------------|---------------------|
| 1 | 63 | 31.3 | 80246 | 3 | US-09-078-294-4 | Sequence 4, Appl1 |
| 2 | 61.4 | 30.5 | 80595 | 3 | US-09-078-294-3 | Sequence 3, Appl1 |
| 3 | 59.8 | 29.8 | 223 | 1 | US-08-222-177A-14 | Sequence 14, Appl1 |
| 4 | 59.6 | 29.7 | 50000 | 4 | US-09-146-053-4 | Sequence 4, Appl1 |
| 5 | 59.2 | 29.5 | 965 | 3 | US-09-280-799-133 | Sequence 133, Appl1 |
| 6 | 59 | 29.4 | 538 | 4 | US-09-499-050A-105 | Sequence 105, Appl1 |
| 7 | 58.4 | 29.1 | 2235 | 3 | US-09-334-601-3 | Sequence 3, Appl1 |
| 8 | 58.2 | 29.0 | 747 | 4 | US-09-220-132-92 | Sequence 92, Appl1 |
| 9 | 58.2 | 29.0 | 35100 | 1 | US-08-306-631B-19 | Sequence 19, Appl1 |
| 10 | 58.2 | 29.0 | 35100 | 5 | PCT-US93-06251-19 | Sequence 19, Appl1 |
| 11 | 58.2 | 29.0 | 48484 | US-09-435-739-42 | Sequence 42, Appl1 | |
| 12 | 57.8 | 28.8 | 2121 | 4 | US-09-425-488-1 | Sequence 1, Appl1 |
| 13 | 57.6 | 28.7 | 4157 | 2 | US-08-871-266B-1 | Sequence 1, Appl1 |
| 14 | 57.6 | 28.7 | 4157 | 2 | US-08-819-458A-1 | Sequence 1, Appl1 |
| 15 | 57.6 | 28.7 | 4157 | 2 | US-09-018-664A-1 | Sequence 1, Appl1 |
| 16 | 57.6 | 28.7 | 4157 | 3 | US-08-871-267B-1 | Sequence 1, Appl1 |
| 17 | 57.6 | 28.7 | 4157 | 3 | US-08-618-419-1 | Sequence 1, Appl1 |
| 18 | 57.6 | 28.7 | 8758 | 4 | US-09-799-345-3 | Sequence 3, Appl1 |
| 19 | 57.6 | 28.7 | 8758 | 4 | US-09-963-276-3 | Sequence 3, Appl1 |
| 20 | 57.2 | 28.5 | 63588 | 4 | US-09-897-404-3 | Sequence 3, Appl1 |
| 21 | 57 | 28.4 | 3604 | 4 | US-09-688-158B-27 | Sequence 27, Appl1 |
| 22 | 57 | 28.4 | 3604 | 4 | US-09-291-417D-27 | Sequence 27, Appl1 |
| 23 | 57 | 28.4 | 10409 | 3 | US-08-772-440-33 | Sequence 33, Appl1 |
| 24 | 57 | 28.4 | 55298 | 4 | US-09-491-356C-1 | Sequence 1, Appl1 |
| 25 | 57 | 28.4 | 118057 | 4 | US-09-497-855A-32 | Sequence 32, Appl1 |
| 26 | 56.8 | 28.3 | 1920 | 4 | US-09-534-638-6 | Sequence 6, Appl1 |
| 27 | 56.8 | 28.3 | 9840 | 4 | US-09-534-638-6 | Sequence 1, Appl1 |

| | | | | | | | |
|---|----|------|------|--------|---|--------------------|-------------------|
| C | 28 | 56.6 | 28.2 | 203.6 | 4 | US-09-023.655-921 | Sequence 921, App |
| C | 29 | 56.6 | 28.2 | 259.9 | 4 | US-09-051.969A-9 | Sequence 5 |
| C | 30 | 56.4 | 28.1 | 347.4 | 4 | US-09-027.236A-1 | Sequence 1, Appl1 |
| C | 31 | 56.4 | 28.1 | 347.4 | 4 | US-09-056.854-1 | Sequence 1, Appl1 |
| C | 32 | 56.4 | 28.1 | 3996.2 | 4 | US-09-020.924-3 | Sequence 3, Appl1 |
| C | 33 | 56.2 | 28.0 | 202 | 4 | US-09-018.930.115 | Sequence 115, App |
| C | 34 | 56.2 | 28.0 | 202 | 4 | US-09-012.283C-1 | Sequence 115, App |
| C | 35 | 56 | 27.9 | 57 | 4 | US-08-022.177A-362 | Sequence 382, App |
| C | 36 | 56 | 27.9 | 328.8 | 2 | US-08-020.694A-1 | Sequence 1, Appl1 |
| C | 37 | 56 | 27.9 | 328.8 | 3 | US-09-022.255-1 | Sequence 1, Appl1 |
| C | 38 | 56 | 27.9 | 328.8 | 3 | US-09-022.696A-1 | Sequence 1, Appl1 |
| C | 39 | 56 | 27.9 | 328.8 | 3 | US-08-978.777-1 | Sequence 1, Appl1 |
| C | 40 | 56 | 27.9 | 328.8 | 3 | US-09-022.253-1 | Sequence 1, Appl1 |
| C | 41 | 56 | 27.9 | 328.8 | 3 | US-09-022.260-1 | Sequence 1, Appl1 |
| C | 42 | 56 | 27.9 | 328.8 | 3 | US-09-022.259-1 | Sequence 1, Appl1 |
| C | 43 | 56 | 27.9 | 328.8 | 3 | US-09-022.257-1 | Sequence 1, Appl1 |
| C | 44 | 56 | 27.9 | 328.8 | 4 | US-09-042.679-1 | Sequence 1, Appl1 |
| C | 45 | 55.8 | 27.8 | 3000 | 1 | US-08-037.985-3 | Sequence 3, Appl1 |

ALIGNMENTS

```

RESULT 1
US-09-078-294-4
; Sequence 4, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 80246
; TYPE: DNA
; ORGANISM: Nucleotide sequence of NC-contig
; US-09-078-294-4

```

| | | | | |
|-----------------------|-----------------|--------------------|-----------|---------------|
| Query Match | 31.3%; | Score 63; | DB 3; | Length 80246; |
| Best Local Similarity | 87.3%; | Pred. No. 1.5e-09; | | |
| Matches 69; | Conservative 0; | Mismatches 10; | Indels 0; | Gaps 0; |

[illegible]

ORGANISM: Nucleotide sequence of HC-contig
US-09-078-294-3

Query Match 30.5%; Score 61.4; DB 3; Length 80595;
Best Local Similarity 86.1%; Pred. No. 4.7e-09;
Matches 68; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 120 CTCTCTCTCTCCACACACACACACACACACACACACACACACACAA 179
DB 52316 CTCTCTCTCTGACACACACACACACACACACACACACACACACAC 52375
QY 180 ACACACACCCCGCCACCA 198
DB 52376 ACACACACACACACACACACAC 52394

RESULT 3

US-08-222-177A-14
Sequence 14, Application US/08222177A

Patent No. 5582979

GENERAL INFORMATION:

APPLICANT: Weber, James L.

TITLE OF INVENTION: LENGTH POLYMORPHISMS IN

TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME

NUMBER OF SEQUENCES: 460

CORRESPONDENCE ADDRESS:

ADDRESSEE: Demiltt Ross & Stevens, S.C.

STREET: 8000 Excelsior Drive, Suite 401

CITY: Madison

STATE: Wisconsin

COUNTRY: USA

ZIP: 53717-1914

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/222.177A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/341,562

FILING DATE: 21-APR-1989

ATTORNEY/AGENT INFORMATION:

NAME: Sara, Charles S.

REGISTRATION NUMBER: 30,492

REFERENCE/DOCKET NUMBER: 09865, 601

TELEPHONE: (608) 831-2100

TELEFAX: (608) 831-2106

TELEX:

INFORMATION FOR SEQ. ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 223 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

INDIVIDUAL ISOLATE: Caucasian

TISSUE TYPE: Blood

IMMEDIATE SOURCE:

CLONE: MT426

POSITION IN GENOME:

CHROMOSOME/SEGMENT: 18

FEATURE:

NAME/KEY: repeat region

LOCATION: 129..185

OTHER INFORMATION: /rpt_type="candem"

OTHER INFORMATION: /rpt_family="(dc-da)n.(dg-dt)n"

OTHER INFORMATION: /citation="(12)"

FEATURE:

NAME/KEY: misc feature

LOCATION: 90..109

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /evidence= EXPERIMENTAL

OTHER INFORMATION: /standard_name="PCR primer"

OTHER INFORMATION: /citation="(11)"

FEATURE:

NAME/KEY: misc feature

LOCATION: complement (185..204)

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /evidence= EXPERIMENTAL

OTHER INFORMATION: /standard_name="PCR primer"

OTHER INFORMATION: /citation="(11)"

FEATURE:

NAME/KEY: misc feature

LOCATION: 1..223

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /evidence= EXPERIMENTAL

OTHER INFORMATION: /standard_name="Only one strand sequenced"

OTHER INFORMATION: /citation="(11)"

FEATURE:

NAME/KEY: misc feature

LOCATION: 18

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /evidence= EXPERIMENTAL

OTHER INFORMATION: /standard_name="Only one strand sequenced"

OTHER INFORMATION: /citation="(11)"

FEATURE:

NAME/KEY: misc feature

LOCATION: 18

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /evidence= EXPERIMENTAL

OTHER INFORMATION: /standard_name="Only one strand sequenced"

OTHER INFORMATION: /citation="(11)"

FEATURE:

NAME/KEY: misc feature

LOCATION: 18

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /evidence= EXPERIMENTAL

OTHER INFORMATION: /standard_name="Only one strand sequenced"

OTHER INFORMATION: /citation="(11)"

FEATURE:

NAME/KEY: misc feature

LOCATION: 18

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /evidence= EXPERIMENTAL

OTHER INFORMATION: /standard_name="Only one strand sequenced"

OTHER INFORMATION: /citation="(11)"

FEATURE:

NAME/KEY: misc feature

LOCATION: 18

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /evidence= EXPERIMENTAL

OTHER INFORMATION: /standard_name="Only one strand sequenced"

OTHER INFORMATION: /citation="(11)"

FEATURE:

NAME/KEY: misc feature

LOCATION: 18

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /evidence= EXPERIMENTAL

OTHER INFORMATION: /standard_name="Only one strand sequenced"

OTHER INFORMATION: /citation="(11)"

FEATURE:

NAME/KEY: misc feature

LOCATION: 18

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /evidence= EXPERIMENTAL

OTHER INFORMATION: /standard_name="Only one strand sequenced"

OTHER INFORMATION: /citation="(11)"

RESULT 4

US-09-146-053-4/C
Sequence 4, Application US/09146053A

Patent No. 6399349

GENERAL INFORMATION:

APPLICANT: Ryan, James W.

APPLICANT: Sprinkle, Terry Joe Curtis

APPLICANT: Venema, Richard C.

TITLE OF INVENTION: Human Aminopeptidase P Gene

FILE REFERENCE: MCG103

CURRENT APPLICATION NUMBER: US/09/146,053A

CURRENT FILING DATE: 1998-09-02

EARLIER APPLICATION NUMBER: 60/057,854

EARLIER FILING DATE: 1997-09-02

NUMBER OF SEQ ID NOS: 7

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 4

LENGTH: 50000

TYPE: DNA

Query Match

Best Local Similarity 29.8%; Score 59.8; DB 1; Length 223;

Matches 79; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 89 TCACGACACTTTTCTTCTTATGATAGGCTCTCTCTCCACACACACACAA 148
DB 88 TTCAGAAATTTCTCTGCGTATTTTCAATTCTTTTGAACACACACACACAA 147
QY 149 CAC 199
DB 148 CAC 198

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2004, 17:19:49 ; Search time 352.627 Seconds
(without alignments)
2098.537 Million cell updates/sec

Title: US-09-966-880A-9_COPY_2707_2907
Perfect score: 201
Sequence: 1 gtagagcttattatgatttt.....acacaccgcgcacacaag 201

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEM_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 201 | 100.0 | 5514 | 9 | US-09-966-880A-9 |
| 2 | 201 | 100.0 | 11204 | 9 | US-09-966-880A-35 |
| 3 | 70 | 34.8 | 3576 | 11 | US-09-245-277-37 |
| 4 | 70 | 34.8 | 3578 | 11 | US-09-244-805-37 |
| 5 | 69.4 | 34.5 | 6927 | 15 | US-10-423-681-4 |
| 6 | 67.2 | 33.4 | 647 | 15 | US-10-027-632-275968 |
| 7 | 66.4 | 33.0 | 165961 | 15 | US-10-085-117-241 |
| 8 | 65.2 | 32.4 | 15306 | 14 | US-10-238-676-162 |
| 9 | 65.2 | 32.4 | 15306 | 14 | US-10-240-453-184 |
| 10 | 64.8 | 32.2 | 406 | 15 | US-09-233-785-256 |
| 11 | 64 | 31.8 | 525 | 15 | US-10-027-632-45364 |
| 12 | 64 | 31.8 | 605 | 15 | US-10-027-632-275968 |
| 13 | 64 | 31.8 | 630 | 15 | US-10-027-632-301462 |
| 14 | 63.8 | 31.7 | 362 | 15 | US-10-027-632-147955 |
| 15 | 63.8 | 31.7 | 719 | 15 | US-10-027-632-250183 |

| | | | | | | |
|----|------|------|---------|----|----------------------|--------------------|
| 16 | 63.6 | 31.6 | 462 | 15 | US-10-027-632-275967 | Sequence 275967, A |
| 17 | 63.6 | 31.6 | 468 | 10 | US-09-918-995-14728 | Sequence 14728, A |
| 18 | 63.6 | 31.6 | 5179 | 14 | US-10-311-455-1969 | Sequence 1969, Ap |
| 19 | 63.6 | 31.6 | 5179 | 14 | US-10-240-453-277 | Sequence 277, App |
| 20 | 63.2 | 31.4 | 82938 | 10 | US-09-818-657-3 | Sequence 3, Appli |
| 21 | 63 | 31.3 | 80246 | 10 | US-09-728-552-4 | Sequence 4, Appli |
| 22 | 63 | 31.3 | 263744 | 14 | US-10-229-834A-6 | Sequence 6, Appli |
| 23 | 62.8 | 31.2 | 878 | 15 | US-10-027-632-119922 | Sequence 119922, A |
| 24 | 62.6 | 31.1 | 362 | 15 | US-10-027-632-20442 | Sequence 20442, A |
| 25 | 62.6 | 31.1 | 215980 | 10 | US-09-972-546-16 | Sequence 16, Appli |
| 26 | 62.4 | 31.0 | 26042 | 14 | US-10-004-113-31 | Sequence 31, Appli |
| 27 | 62.2 | 30.9 | 887 | 15 | US-10-027-632-154621 | Sequence 154621, A |
| 28 | 62.2 | 30.9 | 26042 | 14 | US-10-027-632-154622 | Sequence 154622, A |
| 29 | 61.8 | 30.7 | 26042 | 14 | US-10-004-113-31 | Sequence 31, Appli |
| 30 | 61.4 | 30.5 | 80595 | 10 | US-09-728-552-3 | Sequence 3, Appli |
| 31 | 60.8 | 30.2 | 481 | 15 | US-10-027-632-298608 | Sequence 298608, A |
| 32 | 60.8 | 30.2 | 96602 | 15 | US-10-085-117-61 | Sequence 61, Appli |
| 33 | 60.6 | 30.1 | 478 | 15 | US-10-027-632-83629 | Sequence 83629, A |
| 34 | 60.6 | 30.1 | 478 | 15 | US-10-027-632-83630 | Sequence 83630, A |
| 35 | 60.6 | 30.1 | 659158 | 9 | US-09-771-208-20 | Sequence 20, Appli |
| 36 | 60.4 | 30.0 | 575 | 15 | US-10-027-632-279794 | Sequence 279794, A |
| 37 | 60.2 | 30.0 | 452 | 9 | US-09-864-761-651 | Sequence 651, App |
| 38 | 60 | 29.9 | 798 | 15 | US-10-027-632-128100 | Sequence 128100, A |
| 39 | 60 | 29.9 | 798 | 15 | US-10-027-632-128101 | Sequence 128101, A |
| 40 | 60 | 29.9 | 1067 | 15 | US-10-027-632-264022 | Sequence 264022, A |
| 41 | 60 | 29.9 | 1067 | 15 | US-10-027-632-264023 | Sequence 264023, A |
| 42 | 60 | 29.9 | 8995 | 9 | US-09-877-935-1 | Sequence 1, Appli |
| 43 | 59.8 | 29.8 | 15698 | 14 | US-10-311-455-2114 | Sequence 2114, Ap |
| 44 | 59.8 | 29.8 | 1691139 | 14 | US-10-067-514-1 | Sequence 1, Appli |
| 45 | 59.8 | 29.8 | 1691139 | 15 | US-10-419-723-1 | Sequence 1, Appli |

ALIGNMENTS

RESULT 1
US-09-966-880A-9
Sequence 9, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Muramatsu, Masamichi
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001 US/09/966, 880A
CURRENT APPLICATION NUMBER: PCT/JP00/01918
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US/09-966, 880A
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 5514
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Intron
LOCATION: (1) ... (1031)
FEATURE:
NAME/KEY: Exon
LOCATION: (1032) ... (1116)
FEATURE:
NAME/KEY: Intron
LOCATION: (1119) ... (5514)
US-09-966-880A-9
Query Match 100.0%; Score 201; DB 9; Length 5514;
Best Local Similarity 100.0%; Pred. No. 1.4e-51;

Matches 91; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 69 ACATCCTTTTACTTACTTCCAGACATTTTCTCTTATGATAGGCTCTCTC 128
Db 3102 ACAACTC 3043

QY 129 TCTCCACC 188
Db 3042 TCTAC 2983

QY 189 CCGCCA 194
Db 2982 CACACA 2977

RESULT 5
US-10-429-681-4/c

Sequence 4, Application US/10429681
Publication No. US20040003425A1
GENERAL INFORMATION:
APPLICANT: Japan as Represented by Director General of Okazaki National Research
APPLICANT: Institutes
TITLE OF INVENTION: Nav2 channel gene-deficient non-human animals
FILE REFERENCE: U2001P059
CURRENT APPLICATION NUMBER: US/10/429,681
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: JP 2000/237320
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: JP 2000/241637
PRIOR FILING DATE: 2000-08-09
PRIOR APPLICATION NUMBER: JP 2001/222263
PRIOR FILING DATE: 2001-07-23
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 6927
TYPE: DNA
ORGANISM: Mus musculus
US-10-429-681-4

Query Match 34.5%; Score 69.4; DB 15; Length 6927;
Best Local Similarity 92.4%; Pred. No. 5,1e-11;
Matches 73; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 120 CTC 179
Db 3038 CTC 2979

QY 180 ACACACACCCCGCAACCA 198
Db 2978 ACACACACCACTACACCA 2960

RESULT 6
US-10-027-632-261170

Sequence 261170, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108627.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 261170
LENGTH: 647
TYPE: DNA
ORGANISM: Human
US-10-027-632-261170

Query Match 33.4%; Score 67.2; DB 15; Length 647;
Best Local Similarity 85.2%; Pred. No. 1e-10;
Matches 75; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 100 TTTTCTCTTATGATAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 159
Db 73 TTGTTTCTCATGATATGCCATGACTCCCTGCACACACACACACACACACAC 132

QY 160 AC 187
Db 133 AC 160

RESULT 7
US-10-085-117-241

Sequence 241, Application US/10085117
Publication No. US200302334A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
FILE REFERENCE: 529452000121
CURRENT APPLICATION NUMBER: US/10/085,117
CURRENT FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 361
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 241
LENGTH: 165961
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: variation
LOCATION: (1)...(165961)
OTHER INFORMATION: n = any nucleotide
US-10-085-117-241

Query Match 33.0%; Score 66.4; DB 15; Length 165961;
Best Local Similarity 73.3%; Pred. No. 1.4e-09;
Matches 85; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 78 TTATCTTACTTCCAGGACATTTTCTTCTTATATAGGCTCTCTCTCTCTCTC 137
Db 2691 TGAATTATCAGTTAGGGCACTGTGTTCTTTCAAGCCTGGGTTTCCAGACAC 2750

QY 138 AC 193
Db 2751 ACATGAC 2806

RESULT 8
US-10-239-676-162/c

Sequence 162, Application US/10239676
Publication No. US2003082609A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PREPENROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation

FILE REFERENCE: 5013.1003
CURRENT APPLICATION NUMBER: US/10/239,676
CURRENT FILING DATE: 2002-09-24
PRIOR APPLICATION NUMBER: PCT/EP01/03968
DE 10019058.8
DE 10019173.8
DE 10032529.7
DE 10043826.1
PRIOR FILING DATE: 2001-04-06
2000-04-06
2000-04-07
2000-06-30
2000-09-01
NUMBER OF SEQ ID NOS: 228
SEQ ID NO 162
LENGTH: 15306
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-162

Query Match 32.4%; Score 65.2; DB 14; Length 15306;
Best Local Similarity 72.0%; Pred. No. 1.3e-09;
Matches 85; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 82 CTAACTTCAGACACTTTTCTCTTATGATAAGCTCTCTCTCTCCACACACAC 141
DB 10351 CTAAATTTTCACCACTATTTCCTCCGCTTAAACAAACATCTAATCACTCAACACAC 10292
QY 142 AC 199
DB 10291 AC 10234

RESULT 9

US-10-240-453-184/c
Sequence 184, Application US/10240453
Publication No. US20030148326A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
TITLE OF INVENTION: With DNA Transcription
FILE REFERENCE: 5013.1009
CURRENT APPLICATION NUMBER: US/10/240,453
CURRENT FILING DATE: 2002-10-02
PRIOR APPLICATION NUMBER: PCT/EP01/03973
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 350
SEQ ID NO 184
LENGTH: 15306
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-453-184

Query Match 32.4%; Score 65.2; DB 14; Length 15306;
Best Local Similarity 72.0%; Pred. No. 1.3e-09;
Matches 85; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 82 CTAACTTCAGACACTTTTCTCTTATGATAAGCTCTCTCTCTCCACACACAC 141
DB 10351 CTAAATTTTCACCACTATTTCCTCCGCTTAAACAAACATCTAATCACTCAACACAC 10292
QY 142 AC 199
DB 10291 AC 10234

RESULT 10

US-09-232-785-256/c
Sequence 256, Application US/09232785
Publication No. US20030049612A1
GENERAL INFORMATION:
APPLICANT: International Paper Co.
APPLICANT: Eght, Craig S
APPLICANT: Nelson, C. Dana
TITLE OF INVENTION: MICROSTATELITE DNA MARKERS AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: 4481/1E18US1
CURRENT APPLICATION NUMBER: US/09/232,785
CURRENT FILING DATE: 1999-01-19
PRIOR APPLICATION NUMBER: 09/232,884
PRIOR FILING DATE: 1999-01-15
NUMBER OF SEQ ID NOS: 397
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 256
LENGTH: 406
TYPE: DNA
ORGANISM: Pinus taeda L.
FEATURE:
NAME/KEY: unsure
LOCATION: n at 43 and 406.
OTHER INFORMATION: n is a or g or c or t/u, unknown or other.
US-09-232-785-256

Query Match 32.2%; Score 64.8; DB 10; Length 406;
Best Local Similarity 90.8%; Pred. No. 4.8e-10;
Matches 69; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 112 TGATAGGCTCTCTCTCTCTCCACACACACACACACACACACACACACACAC 171
DB 357 TGAGATGATCTCTCTCTCTCTATACACACACACACACACACACACACACACAC 298
QY 172 ACACACAAACACACAC 187
DB 297 ACACACACACACACAC 282

RESULT 11

US-10-027-632-45364/c
Sequence 45364, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002

;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 45364
;; LENGTH: 525
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-45364

Query Match 31.8%; Score 64; DB 15; Length 525;
Best Local Similarity 60.6%; Pred. No. 9.2e-10;
Matches 103; Conservative 1; Mismatches 66; Indels 0; Gaps 0;

QY 11 ATGTATTTTCATGTAGTAGAGACCAAAATCCAGCCAAAGTCCTTATCTATGCCAC 70
DB 274 ATGTACTCTCCGTAAAGAAATAGCAATGACTTGTATCAATAGACGGGCAAC 215
QY 71 ATCCTTCTATCTATCTTCCAGACACTTTTCTTCTTATGATTAAGGCTCTCTCTC 130
DB 214 ACAATGCTTTTACATAGTTATTAACCTTTTCTTATCTTATCCCAAGTCTCCACCTG 155
QY 131 TCAA 180
DB 154 TACACACACACACACACACAAACACACACACACACACACACACACACACACACAC 105

RESULT 12
US-10-027-632-275968
;; Sequence 275968, Application US/10027632
;; Publication No. US20030204075A9
;; GENERAL INFORMATION:

;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; PRIOR FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 275968
;; LENGTH: 605
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-275968

Query Match 31.8%; Score 64; DB 15; Length 605;
Best Local Similarity 93.1%; Pred. No. 9.7e-10;
Matches 67; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 118 GGCTCTCTCTCTCCAC 177
DB 392 GTCTCTCTCTCTCCATTAACACACACACACACACACACACACACACACACACAC 451
QY 178 AAACACACACCC 189
DB 452 ACACACACACAC 463

RESULT 13

US-10-027-632-301462/c
;; Sequence 301462, Application US/10027632
;; Publication No. US20030204075A9
;; GENERAL INFORMATION:

;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; PRIOR FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 301462
;; LENGTH: 630
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-301462

Query Match 31.8%; Score 64; DB 15; Length 630;
Best Local Similarity 60.6%; Pred. No. 9.8e-10;
Matches 103; Conservative 1; Mismatches 66; Indels 0; Gaps 0;

QY 11 ATGTATTTTCATGTAGTAGAGACCAAAATCCAGCCAAAGTCCTTATCTATGCCAC 70
DB 292 ATGTACTCTCCGTAAAGAAATAGCAATGACTTGTATCAATAGACGGGCAAC 233
QY 71 ATCCTTCTATCTATCTTCCAGACACTTTTCTTCTTATGATTAAGGCTCTCTCTC 130
DB 232 ACAATGCTTTTACATAGTTATTAACCTTTTCTTATCTTATCCCAAGTCTCCACCTG 173
QY 131 TCAA 180
DB 172 TACACACACACACACACACAAACACACACACACACACACACACACACACACACAC 123

RESULT 14
US-10-027-632-147955/c
;; Sequence 147955, Application US/10027632
;; Publication No. US20030204075A9
;; GENERAL INFORMATION:

;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; PRIOR FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002

Job time : 355.627 secs

/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 147955
/ LENGTH: 362
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-147955

Query Match 31.7%; Score 63.8; DB 15; Length 362;
Best Local Similarity 88.3%; Pred. No. 9.3e-10;
Matches 68; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 116 AAGCTCTCTCTCTCCACACACACACACACACACACACACACACACACACACAC 175
DB 345 AACT 286
QY 176 ACAAAACACACACCCCGC 192
DB 285 ACACACACACATCACAC 269

RESULT 15
US-10-027-632-250183
/ Sequence 250183, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 250183
/ LENGTH: 719
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-250183

Query Match 31.7%; Score 63.8; DB 15; Length 719;
Best Local Similarity 60.8%; Pred. No. 1.2e-09;
Matches 104; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 24 GTAGAGTAGAGACCAAAATCCACCAAGCTTTATCTATGACACATCTTTATCT 83
DB 7 GTAAAGTACTAAGTTAAAAAACTTTAGCCCTCTGTCTCTCCAAAGTTGGTATATTA 66
QY 84 ATACTTCAGACACTTTTCTTCTCTATGATAGGCTCTCTCTCTCTCTCTCTCTCT 143
DB 67 AATATATGATATCATTTTATAGCCTGGGAGCATGTGAGACTGTGTCTCATATACACAC 126
QY 144 AC 194
DB 127 AC 177

44 TATTACTCTTCTTTTGTGAGATGGAGTTTGCGTCCTGTG 103

[illegible]

```

RESULT 2
US-09-818-512-3/c
; Sequence 3, Application US/09818512
; Patent No. 6537780
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01192
; CURRENT APPLICATION NUMBER: US/09/818,512
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 116592
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) .. (116592)
; OTHER INFORMATION: n = A,T,C or G
US-09-818-512-3

```

| | | | | | |
|----|--|--------|--------------------|-----------|----------------|
| | Query Match | 64.5% | Score 194.2; | DB 4; | Length 116592; |
| | Best Local Similarity | 84.4%; | Pred. No. 3.7e-37; | | |
| | Matches 217; Conservative | 1; | Mismatches 39; | Indels 0; | Gaps 0; |
| QY | 42 AGATTACTCTTCTTTTTTTTTTTTTTTTTTTTTTTTTGAGATGAGAGTTTGCTGT | 101 | | | |
| Db | 39805 ATTTCGAAATTTTTTTTTTTTTTTTTTTTTTTTTTTGGATGGAGTACATCTTGT | 39746 | | | |
| QY | 102 TGCCCATGCTGAGTGAATGGCATGAYCATACTCACTGCAACTTCACCTTCCTGGT | 161 | | | |
| Db | 39745 TGCCAGGCTGAGAGCGAGTGCGACAATTCAGACACTCAAAGTCCACCTCATGGCT | 39685 | | | |
| QY | 162 CAAGCAAAAGCTGTGCTCAGCTCCCGGGTAGATGGAGTTACAGGGGCCAACACCACA | 221 | | | |
| Db | 39685 CAAGCAATTCCTCAGCCTCAGCCTCCAGAGCTGGAGTTACAGGACCCCATTAACAA | 39626 | | | |
| QY | 222 CTCGGCTAATGTTGTATTTTAAAGANAGGGGTTCCACANTGTGGCCAGGCTGTC | 281 | | | |
| Db | 39625 CCCAGCTAATTTTGTATTTTAAAGAGACAGGGTTTCGCCAGTGTGGCAGGGTGC | 39566 | | | |
| QY | 282 TCAAACCTCCTGACCTCA | 298 | | | |
| Db | 39565 TCAAACCTCCTGACCTCA | 39549 | | | |

RESULT 3
US-09-791-211-10/c
; Sequence 10, Application US/09791211
; Patent No. 6448080
; GENERAL INFORMATION:

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

| | Query | Match | Score | DB | Length |
|----|-----------------------|--|--------------|----|----------------------------------|
| | Best Local Similarity | 63.7% | 191.8 | 4 | 98844 |
| | Matches | 211 | Conservative | 1 | Mismatches 33; Indels 0; Gaps 0; |
| Qy | 55 | CTTTTTCCTTTTTTTTTTTTTTTTGAGATGAGATTGTGCTGTGGTCCCATCTGGA | 114 | | |
| Db | 82331 | CTTTTTCCTTTTTTTTTTTTTTTTGTGAGACAAAGTTTGCTCTTGTGCCAGGCTGGA | 82272 | | |
| Qy | 115 | GTCGAATGGCAGTAGCATAGCTCACTGCACCTCCACTCTCGGGTTCAAAGAAGCTGT | 174 | | |
| Db | 82271 | GTGGAATGGACAACATCTCGGCTCACTGCACCTCCACTCCGAGTTCAAAGATTTCCCC | 82212 | | |
| Qy | 175 | CGCTCACGCTCCCGGGTAGATGGAGATTACAGCGGCCAACACACACACTCGGCTAATGTT | 234 | | |
| Db | 82211 | TGCTCTACGCTCCCAATAAGCTGGAGATTACAGGCACTCCACCAAGGCCACGTAATGTT | 82152 | | |
| Qy | 235 | TGTAATTTTATAGTAGAGATGGGGTTTACCATGTTGGCCAAGGCTGATCAAACTCTGCAC | 294 | | |
| Db | 82151 | TGTAATTTTATAGTAGACAGGGTTTACCATGTTGGCCAAGGCTGATCTGAACCTCTGCAC | 82092 | | |
| Qy | 295 | CTCAG | 299 | | |
| Db | 82091 | CTCAG | 82087 | | |

```

RESULT 4
US-09-548-797B-7/C
; Sequence 7, Application US/09548797B
; Patent No. 6683165
; GENERAL INFORMATION:
; APPLICANT: KEITH, TIM
; TITLE OF INVENTION: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES AND
; TITLE OF INVENTION: OBSTITY
; FILE REFERENCE: 2976-4039

```

```

:
: CURRENT APPLICATION NUMBER: US-09/548,797B
:
: CURRENT FILING DATE: 2002-11-26
:
: PRIOR APPLICATION NUMBER: 60/129,351
:
: PRIOR FILING DATE: 1999-04-13
:
: NUMBER OF SEQ ID NOS: 170
:
: SOFTWARE: PatentIn Ver. 2.1
:
: SEQ ID NO 7
:
: LENGTH: 17000
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: US-09-548-797B-7

```

| | | | | | | |
|----|---|-----------------|--------------------|-----------|---------------|--|
| | Query Match | 62.7%; | Score 188.8; | DB 4; | Length 17000; | |
| | Best Local Similarity | 81.6%; | Pred. No. 4.3e-36; | | | |
| | Matches 217; | Conservative 1; | Mismatches 48; | Indels 0; | Gaps 0; | |
| Qy | 35 ATGATCAAGTATTACTCTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTGGATGAGATTTC | 94 | | | | |
| Dd | 16118 ATTAAACGTGTAATTTTAGAGCTCTCTTTTTTTTTTTTTTTTTTTTGAGAAGCGACTTC | 16059 | | | | |
| Qy | 95 GTCTTGTTGCCCATGCTGCAGTGAATGAGATGACATGACCTCAGCAACTCCACCTC | 154 | | | | |
| Dd | 16058 CTCTTGTTGCCCAAGCTGAGTGAAGTGAACAGATCAGACTCAGCAAACTCCACCTC | 15939 | | | | |
| Qy | 155 CTGGGTTCAAGCAAAAGCTGTGCCTCAGCCTCCCAGGTAATGGGATTACAGGCGCCAC | 214 | | | | |
| Dd | 15998 CCATGTTCAAAACAATTTCTCTCCGCGCAGCCTCCCGAGTACCTGGGATTATAGCAAGTCC | 15939 | | | | |
| Qy | 215 CACCAACTCGGCGTAATGTTTGTATTTTGTAGTAAAGATGGGGTTTACCATAATGTGGCCAG | 274 | | | | |
| Dd | 15938 CACCAACACCGGCGTAATTTTGTATTTTGTAGTAAAGATGGGGTTTACCACATTTGGCCAG | | | | | |
| Qy | 275 GCTGATCGAACAATCCTGACCTCAGA 300 | | | | | |
| Dd | 15878 GCTGATCTCGAACTTCTGACCTCAAA 15853 | | | | | |

```

/ RESULT 5
/ US-09-608-285A-e
/ Sequence 8, Application US/09608285A
/ Patent No. 6335013
/ GENERAL INFORMATION:
/ APPLICANT: Ford, John
/ APPLICANT: Mulero, Julio
/ APPLICANT: Yeung, George
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
/ TITLE OF INVENTION: POLYPEPTIDES
/ FILE REFERENCE: 28110/36570
/ CURRENT APPLICATION NUMBER: US/09/608,285A
/ CURRENT FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: 09/583,231
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 09/557,800
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/481,238
/ PRIOR FILING DATE: 2000-01-11
/ PRIOR APPLICATION NUMBER: 09/370,265
/ PRIOR FILING DATE: 1999-08-09
/ PRIOR APPLICATION NUMBER: PCT/US99/16180
/ PRIOR FILING DATE: 1999-07-16
/ PRIOR APPLICATION NUMBER: 09/350,836
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: 09/273,447
/ PRIOR FILING DATE: 1999-03-19
/ PRIOR APPLICATION NUMBER: 09/244,444
/ PRIOR FILING DATE: 1999-02-04
/ PRIOR APPLICATION NUMBER: 09/122,449
/ PRIOR FILING DATE: 1998-07-24
/ PRIOR APPLICATION NUMBER: 09/118,205
/ PRIOR FILING DATE: 1998-07-16
/ NUMBER OF SEQ ID NOS: 60
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 8

```

```

? LENGTH: 9365
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (3409)
? OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
? NAME/KEY: misc_feature
? LOCATION: (9214)
? OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
? NAME/KEY: misc_feature
? LOCATION: (9303)
? OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
? NAME/KEY: misc_feature
? LOCATION: (9111)
? OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
US-03-608-265A-8

```

| | Query Match | Similarity | 62.3% | Score 187.4 | DB 4 | Length 9365 | | |
|----|-------------|--|--------------|-----------------|------------|-------------|--------|------|
| | Best Local | Similarity | 75.4% | Pred. No. 8e-36 | | | | |
| | Matches | 212 | Conservative | 23 | Mismatches | 44 | Indels | Gaps |
| QY | 21 | ATTCAGTTGACATGATCAAGTATTACTCTTCTTTTTTTTTTTTTTTTTTTTTTTT | | | | | | 80 |
| DB | 7364 | AATTAAAGTGCCAAAGTCATCAGTGTATTATTATTATTATTATTATTATTATT | | | | | | 7443 |
| QY | 81 | GAGATGAGTTTGATCTTTGGTCCCATGCTGAAGTGAATGCATGACTACT | | | | | | 140 |
| DB | 7444 | GAGAGGAGCTGTGGCTCTTTGTTKCYRRGGCTGAGTGCAPRGGRYGATCWRCGTCACT | | | | | | 7503 |
| QY | 141 | GCAACTCCACTCTCTCGGTTCCAAAGCAAAGCTGTGCGCTCAGCGTCCCCGGTAGTGGSA | | | | | | 200 |
| DB | 7504 | GCARCTTCRCCTCTCGGTTCAAGSCAATCTCTGCTCAGCCCTCCRAGTAGCTGGSA | | | | | | 7563 |
| QY | 201 | TTTCAAGGCCCAACCAACAACAAGTGGGCTAA--TGTTTTATTATTTTAAGTAGAGATGGGGTT | | | | | | 258 |
| DB | 7564 | TTTACAGGGRYCGCCACCACCGCCYGGCTAAATTTTTTTTATTTTAAAGTAGAGACGGGGTT | | | | | | 7622 |
| QY | 259 | TCAACATGTGGCCAGGCTGATTCACAACTCTGACCTCAG | | | | | | 299 |
| DB | 7624 | TCAACATGTGGCCAGGCTTRKTCTTAACTCTGATCTCAG | | | | | | 7664 |

```

RESULT 6
US-09-350-836B-8
Sequence 8, Application US/09350836B
Patent No. 6387645
GENERAL INFORMATION:
APPLICANT: Mulero, John
APPLICANT: Mulero, Julio
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
FILE REFERENCE: 28110/35761
CURRENT APPLICATION NUMBER: US/09/350, 836B
CURRENT FILING DATE: 1999-07-09
PRIORITY APPLICATION NUMBER: 09/273, 447
PRIORITY FILING DATE: 1999-03-19
PRIORITY APPLICATION NUMBER: 09/118, 205
PRIORITY FILING DATE: 1998-07-16
PRIORITY APPLICATION NUMBER: 09/122, 449
PRIORITY FILING DATE: 1998-07-24
PRIORITY APPLICATION NUMBER: 09/244, 444
PRIORITY FILING DATE: 1999-02-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 9365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: exon
LOCATION: (1)..(288)
NAME/KEY: exon

```

| | | | | |
|-----------------------|------------------|------------------|-----------|--------------|
| Query Match | 62.3%; | Score 187.4; | DB 4; | Length 9365; |
| Best Local Similarity | 75.4%; | Pred. No. 8e-36; | | |
| Matches 212; | Conservative 23; | Mismatches 44; | Indels 2; | Gaps 1. |

| | | | |
|----|------|---|------|
| Qy | 21 | AATTCAGTTTGACAAATGACAGATACTACTCTTCTTTTTTTTTTTTTTTTTTTT | 80 |
| Dd | 7384 | AATTAGAATGCCAAGTCACATGATTTATTATTATTATTATTATTATTATTTTTT | 7443 |
| Qy | 81 | GAGATGAATTTGGTGTCTGTTGCCCATGCTGAGTGAATGCGATGATAGCTACT | 140 |
| Dd | 7444 | GAGACGAGATGTTCCTGTCTTCCYRGGCTGGAGTCATGGCGRYATCMCGCTCACT | 7503 |
| Qy | 141 | GCAACTGCACCTCCTGGGTTCAACAAGCTGACGCTCAGCCTCCCGGGTAGATGGA | 200 |
| Dd | 7504 | GCACTCYRCCTCTCGGGTTTAACGATTTCTTCTGCTCAGCCTCCCRAGTACGTGGA | 7563 |
| Qy | 201 | TTCAGCGCCCAACCAACAACACTCGGCTAA--TGTTGTATTTTAAAGATGGGGTT | 258 |
| Dd | 7564 | TTACAGCGRCGCCCAACCAACGCGCCYAGCTAATTTTTTTGTATTTTAAAGACGGGGTT | 7623 |
| Qy | 259 | TCAACATGTTGGCGAGGCTGCTGCTAAATCTGTACCTCGAG | 299 |
| Dd | 7624 | TCAACATGTTGGCGAGGCTKATCTTAACTGTATGATCTCAG | 7664 |

```

RESULT 7
US-09-370-265-8
; Sequence 8, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Jullio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 2811/35908
; CURRENT APPLICATION NUMBER: US/09/370,265
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: PCT/US99/16180
; EARLIER FILING DATE: 1999-07-16
; EARLIER APPLICATION NUMBER: 09/350,836
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: 09/273,447
;

```

```

1 TYPE: DNA
2 ORGANISM: Homo sapiens
3
4 FEATURE:
5
6 NAME/KEY: CDS
7 LOCATION: (72)..(287)
8 FEATURE:
9 NAME/KEY: CDS
10 LOCATION: (1280)..(1579)
11 FEATURE:
12 NAME/KEY: CDS
13 LOCATION: (1819)..(1854)
14 FEATURE:
15 NAME/KEY: CDS
16 LOCATION: (2466)..(2555)
17 FEATURE:
18 NAME/KEY: CDS
19 LOCATION: (2863)..(2940)
20 FEATURE:
21 NAME/KEY: CDS
22 LOCATION: (3887)..(3952)
23 FEATURE:
24 NAME/KEY: CDS
25 LOCATION: (4896)..(4994)
26 FEATURE:
27 NAME/KEY: CDS
28 LOCATION: (5846)..(5986)
29 FEATURE:
30 NAME/KEY: CDS
31 LOCATION: (6965)..(7138)
32 FEATURE:
33 NAME/KEY: CDS
34 LOCATION: (8556)..(8639)
35 FEATURE:
36 NAME/KEY: misc_feature
37 LOCATION: (3403)
38 OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
39 FEATURE:
40 NAME/KEY: misc_feature
41 LOCATION: (9214)
42 OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
43 FEATURE:
44 NAME/KEY: misc_feature
45 LOCATION: (9303)
46 OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
47 FEATURE:
48 NAME/KEY: misc_feature
49 LOCATION: (9311)
50 OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
51
52 JS-09-370-265-8

```

| | | | | |
|-----------------------|--------------|------------------|----------------|--------------|
| Query Match | 62.3%; | Score 187.4; | DB 4; | Length 9365; |
| Best Local Similarity | 75.4%; | Pred. No. 8e-36; | | |
| Matches 212; | Conservative | 23; | Mismatches 44; | Indels 2; |
| | | | | Gaps 1. |

[illegible]

| | | | |
|----|------|--|------|
| OY | 141 | GCACCTCTCACTCCTGGGTTCAAGCAAGCTGTGGCTCAGACCTCCGGGTAAATGGGA | 200 |
| Db | 7504 | GCACCTCTCCTCTCTGGTTTCAAGCATTTCTTCGTCTCAGCTCCCACTGAGTGGGA | 7563 |
| OY | 201 | TTACAGGGCCACCACCACCACTCGGCTAA--TGTTTATTTTATGAGAGATGGGGTT | 258 |
| Db | 7564 | TTACAGGCTTGGGCGCCACCAAGCTGGCTAATTTTTTTGTATTTTATGTAAGACGGGGTT | 7623 |
| OY | 259 | TCACCATTTTGCCAGAGCTGGATCTCAAACTCCTGACCTCAG | 299 |
| Db | 7624 | TCACCATTTTGCCAGAGCTTCTCTTATCTCTGATGCTCAG | 7664 |

RESULT 8
US-09-557-800C-8
; Sequence 8, Application US/09557800C

```

1  APPLICANT: Mulero, Julio
2  APPLICANT: Yeung, George
3  TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
4  TITLE OF INVENTION: Polypeptides
5  FILE REFERENCE: 28110/36457
6  CURRENT APPLICATION NUMBER: US/09/557,800C
7  CURRENT FILING DATE: 2000-04-25
8  PRIOR APPLICATION NUMBER: 09/481,238
9  PRIOR FILING DATE: 2000-01-11
10 PRIOR APPLICATION NUMBER: 09/370,265
11 PRIOR FILING DATE: 1999-08-09
12 PRIOR APPLICATION NUMBER: PCT/US99/16180
13 PRIOR FILING DATE: 1999-07-16
14 PRIOR APPLICATION NUMBER: 09/350836
15 PRIOR FILING DATE: 1998-07-09
16 PRIOR APPLICATION NUMBER: 09/273447
17 PRIOR FILING DATE: 1999-03-19
18 PRIOR APPLICATION NUMBER: 09/122449
19 PRIOR FILING DATE: 1998-07-24
20 PRIOR APPLICATION NUMBER: 09/244444
21 PRIOR FILING DATE: 1998-02-04
22 PRIOR APPLICATION NUMBER: 09/118,205
23 PRIOR FILING DATE: 1998-07-16
24 NUMBER OF SEQ ID NOS: 56
25 SOFTWARE: PatentIn Ver. 2.0
26 SEQ ID NO 8
27 LENGTH: 9365
28 TYPE: DNA
29 ORGANISM: Homo sapiens
30 FEATURE:
31 NAME/KEY: misc feature
32 LOCATION: (3405)
33 OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
34 FEATURE:
35 NAME/KEY: misc feature
36 LOCATION: (9214)
37 OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
38 FEATURE:
39 NAME/KEY: misc feature
40 LOCATION: (9303)
41 OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
42 FEATURE:
43 NAME/KEY: misc feature
44 LOCATION: (9311)
45 OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
46
47 US-09-557-800C-8

```

| | | | | |
|-----------------------|------------------|------------------|-----------|--------------|
| Query Match | 62.3%; | Score 187.4; | DB 4; | Length 9365; |
| Best Local Similarity | 75.4%; | Pred. No. 8e-36; | | |
| Matches 212; | Conservative 23; | Mismatches 44; | Indels 2; | Gaps 1; |

| QY | 22 | ATTCAAGTTGCACATGATCAGTAACTACCTTCGTTTTTTTTTTTTTTTTTT | 80 |
|----|------|---|------|
| Db | 7384 | AATTAGATGCCAAGTCACATGTTTATTATTATTATTATTATTATTATTATTTTTT | 7443 |

| QY | 81 | GAGATGGAGTTTGGCTCTGTGTGGCAGTGGAGTGGATGGACATGACATGCTACT | 140 |
|----|------|---|------|
| Db | 7444 | GAGACGAGATCTGCTCTTGTATTCCTCTGAGCTGAGATGGACATGGCATGCMCRGCTCACT | 7500 |
| QY | 141 | GCAACCTCCACCTCCTGGGTTCAAGCAAGCTGTGCTCAGCTCCCGGTTAGATGGGA | 200 |
| Db | 7504 | GCARCTTCRCTCCTGGGTTCAAGCATTCCTCTGCTCAGGCTCCCAAGTAGCTGGGA | 7563 |
| QY | 201 | TTACAGCGGCCACACACACACTGGGTTA--TGTTTGATTTTGTAGATATGGGTT | 258 |
| Db | 7564 | TTACAGCGCGGCCACACACCGCTGGGTTATTTTGTATTTTGTAGTGAAGCGGGTT | 7623 |
| QY | 259 | TCACCATGTGGCGAGGCTGGTCTCAAACTCTGACCTCAG | 299 |
| Db | 7624 | TCACCATGTGGCGAGGCTRKTCTTAACTCTGATCTCAG | 7664 |

QY 259 TCACCATGTTGGCCAGGTGGTCTCAAACTCCTGACCCTCG 259
| | | | | : : : : :
Db 7624 TCACCATGTTGGCCAGGCTRKTCTYRAACTYCTGAYCTCAG 7664

```

1 RESULT 9
2 US-09-370-625A-8
3 Sequence 8, Application US/09370625A
4 Patent No. 6600032
5 GENERAL INFORMATION:
6 APPLICANT: Ford, John
7 APPLICANT: Mulero, Julio
8 TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
9 FILE REFERENCE: 28110/35908
10 CURRENT APPLICATION NUMBER: US/09/370,625A
11 CURRENT FILING DATE: 1999-08-09
12 PRIOR APPLICATION NUMBER: PCT/US99/16180
13 PRIOR FILING DATE: 1999-07-16
14 PRIOR APPLICATION NUMBER: 09/350,836
15 PRIOR FILING DATE: 1999-07-09
16 PRIOR APPLICATION NUMBER: 09/273,447
17 PRIOR FILING DATE: 1999-03-19
18 NUMBER OF SEQ ID NOS: 39
19 SOFTWARE: PatentIn Ver. 2.0
20 SEQ ID NO 8
21 LENGTH: 9365
22 TYPE: DNA
23 ORGANISM: Homo sapiens
24 FEATURE:
25 NAME/KEY: exon
26 LOCATION: (1)..(288)
27 FEATURE:
28 NAME/KEY: exon
29 LOCATION: (1281)..(1580)
30 FEATURE:
31 NAME/KEY: exon
32 LOCATION: (1820)..(1855)
33 FEATURE:
34 NAME/KEY: exon
35 LOCATION: (2467)..(2555)
36 FEATURE:
37 NAME/KEY: exon
38 LOCATION: (2863)..(2942)
39 FEATURE:
40 NAME/KEY: exon
41 LOCATION: (3889)..(3950)
42 FEATURE:
43 NAME/KEY: exon
44 LOCATION: (4894)..(4995)
45 FEATURE:
46 NAME/KEY: exon
47 LOCATION: (5847)..(5987)
48 FEATURE:
49 NAME/KEY: exon
50 LOCATION: (6966)..(7138)
51 FEATURE:
52 NAME/KEY: exon
53 LOCATION: (8556)..(9365)
54 FEATURE:
55 NAME/KEY: misc_feature

```

```

? LOCATION: (3409)
? OTHER INFORMATION: n = a or g or t or c
?
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (9214)
?
? OTHER INFORMATION: n = a or g or t or c
?
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (9303)
?
? OTHER INFORMATION: n = a or g or t or c
?
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (9311)
?
? OTHER INFORMATION: n = a or g or t or c
?
US-09-370-625A-8

```

| | | | | |
|-----------------------|------------------|------------------|-----------|--------------|
| Query Match | 62.3% | Score 187.4; | DB 4; | Length 9365; |
| Best Local Similarity | 75.4% | Pred. No. 8e-36; | | |
| Matches 212; | Conservative 23; | Mismatches 44; | Indels 2; | Gaps 1; |

| | | | |
|----|------|--|------|
| OY | 21 | AATCAAGTTGCACANAGNCAAGTAATCTCTTCTCTTTTTTTTTTTTTTTTTTTT | 80 |
| Dd | 7364 | AATTAGATGCCAAGTCACATGGTTATTATTATTATTATTATTATTATTATT | 7443 |
| OY | 81 | GAGATGAGTTTGCTTTGTGGCCAGCTGGAGTGGANCGCATGAATAGTCTACT | 140 |
| Dd | 7444 | GAGACGGAGTCTGCTCTTGTTKCYRGGCTGAGTGCARTGGCVRGATGWRGGTCACT | 7503 |
| OY | 141 | GGAACCTCACACTCCTGGGTTCAACGAAGAAGCTGGCCACGCTCCGGGTAGATGGAA | 200 |
| Dd | 7504 | GGARCTTCRCCTCTCTGGGTTCAAGGAGATTCTCTGCTCAGCCTCCCRAGTAGTGGA | 7566 |
| OY | 201 | TTACAGCGCCCCACCACAACAACACTCGGCTAA--TGGTTGATTTTAAAGATGGGGTT | 258 |
| Dd | 7564 | TACAGGGRGCGCCACCAACGCGCGGCTAATTTTGTGATTTTAAAGATAGACGGGGTT | 7623 |
| OY | 259 | TCACCATGTTGGCCAGGCTGCTCAAATCCTGACCTGAG | 299 |
| Dd | 7624 | TCACCATGTTGGCCAGGCTRKCTCTPAACTCTTCGATCTAG | 7664 |

```

RESULT 10
US-09-608-285A-42
; Sequence 42, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205

```

```

? PRIOR FILING DATE: 1998-07-16
? NUMBER OF SEQ. ID NOS: 60
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 42
? LENGTH: 14747
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc feature
? LOCATION: (13641)
? OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine
US-03-608-285A-42

```

| | | | | |
|-----------------------|------------------|------------------|-----------|---------------|
| Query Match | 62.3%; | Score 187.4; | DB 4; | Length 14747; |
| Best Local Similarity | 75.4%; | Pred. No. 9e-36; | | |
| Matches 212; | Conservative 23; | Mismatches 44; | Indels 2; | Gaps 1; |

| | | | |
|----|-------|---|--------|
| CY | 21 | AATCAAGTTGACAACGATCAAGTATTACCTCTCTTTTTTTTTTTTTTTTTTTTTTTT | 80 |
| Db | 10744 | AATTAAATGCCAAGCACAATGTTTTATTATTATTATTATTATTATTATTATTATT | 10803 |
| CY | 81 | GAGATGAGATTGGTTGGTTGTGCCCATGCTGAGTGGAATGGCATGAYCATAGCTCACT | 140 |
| Db | 10804 | GAGCGAGAGTCVTGCCTTTGTTCACVRGGCTGAGTAGCATGRCRGAATCMORGGCTCACT | 10863 |
| CY | 141 | GCAACCTCCACACTCCTGGGTTCAAAGAAAAGCTGTGCCTCAGCGCTCCCGGATGATGGGA | 200 |
| Db | 10864 | GCAACTCYCCCTCTCTGGGTTCAAGGAGATCTCTGCTCAAGCTCCCAAGTAGCTGGGA | 109223 |
| CY | 201 | TTACAGGCGGCCAACCAACACACTCGGCTAA - TGTTTGATTTTTAGTAGAGATGGGGTT | 258 |
| Db | 10924 | TTACAGGCRVGGCGCACACAGCGCVGGCTAATTTTTTTTGATTTTTAGTAGAGCGGGGTT | 10983 |
| CY | 259 | TCAACAGTTTGCCAGGCGTGTCTCAAACTCGACCTGAG 299 | |
| Db | 10984 | TCACCAGTTTGCCAGGCGTRKCTCTTAACACTCTGAATCTCAG 11024 | |

```

1      RESULT 11
2      US-09-557-800C-42
3      ; Sequence 42, Application US/09557800C
4      ; Patent No. 6476211
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Ford, John
7      ; APPLICANT: Mulero, Julio
8      ; APPLICANT: Yeung, George
9      ; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
10     ; TITLE OF INVENTION: Polypeptides
11     ; FILE REFERENCE: 28110/36457
12     ; CURRENT APPLICATION NUMBER: US/09/557,800C
13     ; CURRENT FILING DATE: 2000-04-25
14     ; PRIOR APPLICATION NUMBER: 09/481,238
15     ; PRIOR FILING DATE: 2000-01-11
16     ; PRIOR APPLICATION NUMBER: 09/370,265
17     ; PRIOR FILING DATE: 1999-08-09
18     ; PRIOR APPLICATION NUMBER: PCT/US99/16180
19     ; PRIOR FILING DATE: 1999-07-16
20     ; PRIOR APPLICATION NUMBER: 09/350836
21     ; PRIOR FILING DATE: 1999-07-09
22     ; PRIOR APPLICATION NUMBER: 09/273447
23     ; PRIOR FILING DATE: 1999-03-19
24     ; PRIOR APPLICATION NUMBER: 09/122449
25     ; PRIOR FILING DATE: 1998-07-24
26     ; PRIOR APPLICATION NUMBER: 09/244444
27     ; PRIOR FILING DATE: 1999-02-04
28     ; PRIOR APPLICATION NUMBER: 09/118,205
29     ; PRIOR FILING DATE: 1998-07-16
30     ; NUMBER OF SEQ ID NOS: 56
31     ; SOFTWARE: PatentIn Ver. 2.0
32     ; SEQ ID NO 42
33     ; LENGTH: 14747
34     ; TYPE: DNA
35     ; ORGANISM: Homo sapiens

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2004, 17:19:49 ; Search time 528.064 Seconds
(without alignments)
2098.537 Million cell updates/sec

Title: US-09-966-880A-9_COPY_5214_5514
Perfect score: 301
Sequence: 1 cctctctctctctccatagtcgt.....tcaactcctcctcagag 301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 300.6 | 99.9 | 5514 | 9 | US-09-966-880A-9 |
| 2 | 300.6 | 99.9 | 11204 | 9 | US-09-966-880A-35 |
| 3 | 204 | 67.8 | 26657 | 9 | US-09-810-673A-3 |
| 4 | 199.2 | 66.2 | 626 | 15 | US-10-395-242-3 |
| 5 | 199.2 | 66.2 | 626 | 15 | US-10-027-632-133310 |
| 6 | 199.2 | 66.2 | 626 | 15 | US-09-952-213D-6 |
| 7 | 198.8 | 65.0 | 626 | 15 | US-10-027-632-133311 |
| 8 | 196.4 | 65.2 | 15173 | 9 | US-09-764-869-1981 |
| 9 | 196.4 | 65.2 | 15173 | 14 | US-10-091-504-1981 |
| 10 | 196.4 | 65.2 | 15173 | 15 | US-10-227-577-1981 |
| 11 | 196.4 | 65.2 | 42595 | 11 | US-09-984-429-469 |
| 12 | 196.4 | 65.1 | 2460 | 15 | US-10-108-260A-595 |
| 13 | 195.8 | 65.0 | 777 | 15 | US-10-027-632-172403 |
| 14 | 195.8 | 65.0 | 557 | 15 | US-10-027-632-135991 |
| 15 | 194.8 | 64.7 | 6254 | 9 | US-09-764-868-1495 |

| | | | | | | |
|----|-------|------|--------|----|----------------------|--------------------|
| 16 | 194.2 | 64.5 | 116592 | 9 | US-09-818-512-3 | Sequence 3, Appl1 |
| 17 | 194.2 | 64.5 | 116592 | 14 | US-10-354-065-3 | Sequence 3, Appl1 |
| 18 | 194.2 | 64.5 | 164875 | 15 | US-10-085-117-322 | Sequence 332, App |
| 19 | 193.6 | 64.3 | 2136 | 15 | US-10-027-632-110040 | Sequence 110040, |
| 20 | 193.2 | 64.2 | 9564 | 15 | US-10-027-632-250166 | Sequence 250166, |
| 21 | 193.2 | 64.2 | 9564 | 10 | US-09-764-872-481 | Sequence 481, App |
| 22 | 193.2 | 64.2 | 9557 | 10 | US-09-764-872-479 | Sequence 479, App |
| 23 | 193.2 | 64.1 | 145831 | 9 | US-09-969-708-79 | Sequence 79, Appl |
| 24 | 193.2 | 64.1 | 145831 | 9 | US-09-954-456-2116 | Sequence 2116, Ap |
| 25 | 193.2 | 64.1 | 145831 | 10 | US-09-873-357C-646 | Sequence 646, App |
| 26 | 193.2 | 64.1 | 145831 | 12 | US-10-240-425-363 | Sequence 363, App |
| 27 | 193.2 | 64.1 | 358246 | 15 | US-10-292-798-1095 | Sequence 1095, Ap |
| 28 | 192.6 | 64.0 | 5781 | 9 | US-09-764-868-1374 | Sequence 1374, Ap |
| 29 | 192.6 | 64.0 | 5782 | 9 | US-09-764-868-1375 | Sequence 1375, Ap |
| 30 | 192.4 | 63.9 | 23934 | 9 | US-09-764-860-777 | Sequence 777, App |
| 31 | 192.4 | 63.9 | 23934 | 9 | US-09-764-877-2536 | Sequence 2536, App |
| 32 | 192.4 | 63.9 | 23934 | 9 | US-09-764-877-2544 | Sequence 2544, Ap |
| 33 | 192.4 | 63.9 | 23934 | 10 | US-09-764-891-7210 | Sequence 7210, Ap |
| 34 | 192.4 | 63.9 | 23934 | 14 | US-10-074-095-777 | Sequence 777, App |
| 35 | 192.4 | 63.9 | 23934 | 15 | US-10-212-892-2116 | Sequence 2116, App |
| 36 | 192.4 | 63.9 | 23934 | 15 | US-10-242-515-2536 | Sequence 2536, App |
| 37 | 192.4 | 63.9 | 23934 | 15 | US-10-242-515-2544 | Sequence 2544, Ap |
| 38 | 192.4 | 63.8 | 681 | 15 | US-10-027-632-137222 | Sequence 137222, |
| 39 | 192.4 | 63.8 | 23130 | 9 | US-09-764-869-1892 | Sequence 1892, Ap |
| 40 | 192.4 | 63.8 | 23130 | 14 | US-10-091-504-1892 | Sequence 1892, Ap |
| 41 | 192.4 | 63.8 | 23130 | 15 | US-10-227-577-1892 | Sequence 1892, Ap |
| 42 | 191.4 | 63.6 | 5797 | 10 | US-09-764-891-6093 | Sequence 6093, App |
| 43 | 191.2 | 63.5 | 1035 | 14 | US-10-196-846-4227 | Sequence 4227, App |
| 44 | 191.2 | 63.5 | 3150 | 9 | US-09-764-870-622 | Sequence 622, App |
| 45 | 191.2 | 63.5 | 3150 | 14 | US-10-125-540-622 | Sequence 622, App |

ALIGNMENTS

RESULT 1
US-09-966-880A-9
Sequence 9, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966,880A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 5514
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Intron
LOCATION: (1)... (1031)
FEATURE:
NAME/KEY: exon
LOCATION: (1032)... (1118)
FEATURE:
NAME/KEY: Intron
LOCATION: (1119)... (5514)
US-09-966-880A-9
Query Match 99.9%; Score 300.6; DB 9; Length 5514;
Best Local Similarity 100.0%; Pred. No. 5.8e-61;

| | Matches | 301, | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
|----|---------|--|--------------|----|------------|----|--------|----|------|----|
| QY | 1 | CTCTTCAATTTTCCATAGATATCAAGTTTGACATGATCAAGTATTTACTCTTCTTTT | 60 | | | | | | | |
| Db | 5214 | CTCTTCATTTTTCATAGTATTTCAAGTTTACATGATCAAGTATTTCTTTCTTTT | 5273 | | | | | | | |
| QY | 61 | TTTTTTTTTTTTTTTTTGTAGAGATGAGATTTGGTCTGTGGCCATGCTGAGTGGAA | 120 | | | | | | | |
| Db | 5274 | TTTTTTTTTTTTTTTTTGTAGAGATGAGATTTGGTCTGTGTGGCCATGCTGAGTGGAA | 5333 | | | | | | | |
| QY | 121 | TGGCATGATCATAGCTCACTGCAACCTCCACTCTGGGTTCAAGCAAGCTGTGGCCTC | 180 | | | | | | | |
| Db | 5334 | TGGCATGATCATAGCTCACTGCAACCTCCACTCTGGGTTCAAGCAAGCTGTGGCCTC | 5393 | | | | | | | |
| QY | 181 | AGCCTCCCGGGTAGATGGAGATTACAGGCGGCCACACACACTCGGCTTAATGTTTGATTT | 240 | | | | | | | |
| Db | 5394 | AGCCTCCCGGGTAGATGGAGATTACAGGCGGCCACACACACTCGGCTTAATGTTTGATTT | 5453 | | | | | | | |
| QY | 241 | TTTAAAGATGAGATGGGGTTTCAACCATGTTGGCCAGAGCTGGTCCAAACTCTGAACTCAGA | 300 | | | | | | | |
| Db | 5454 | TTTAAAGATGAGATGGGGTTTCAACCATGTTGGCCAGAGCTGGTCCAAACTCTGAACTCAGA | 5513 | | | | | | | |
| QY | 301 | G 301 | | | | | | | | |
| Db | 5514 | G 5514 | | | | | | | | |

```

RESULT 2
US-09-966-880A-35
; Sequence 35, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966, 880A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 11204
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-966-880A-35

```

| | | | | |
|---------------------------|--------|--------------------|-----------|---------------|
| Query Match | 99.9%; | Score 300.6; | DB 9; | Length 11204; |
| Best Local Similarity | 99.7%; | Pred. No. 7.4e-61; | | |
| Matches 300; Conservative | 1; | Mismatches 0; | Indels 0; | Gaps 0; |

| | | | |
|----|------|--|------|
| Qy | 1 | CTCTCTAATTTTCCATAGTATTCAGTTGACANAGATCAAGATTTCTCTTTT | 60 |
| Db | 4624 | CTCTCTAATTTTCCATAGTATTCAGTTGACANAGATCAAGATTTCTTTT | 4683 |
| Qy | 61 | TTTTTTTTTTTTTTTTTGAGATGSAATTTTGCTTTGTGCCCATGCTGGAGTGA | 120 |
| Db | 4684 | TTTTTTTTTTTTTTTTTTTGAGATGSAATTTTGCTTTGTGCCCATGCTGGAG | 4743 |
| Qy | 121 | TGGCATGACATAGCTCACTGCACCTCCCTGAGGTTCAAGCAAGCTGTGGCTC | 180 |
| Db | 4744 | TGGCATGACCATAGCTCACTGCACCTCCCTGAGGTTCAAGCAAGCTGTGGCTC | 4803 |
| Qy | 181 | AGCTCCCGGGTAGATGGGATTACAGGCGCCACACACACTGGGCTAATTTGTATT | 240 |
| Db | 4804 | AGCTCCCGGGTAGATGGGATTACAGGCGCCACACACACTGGGCTAATTTGTATT | 4866 |

QY 241 TTGTGTGAGATGGGGTTTCACCATGTGGGCGAGCTGGTCTCAAACTCCGACCTCAGA 300
DB 4864 TTGTGTGAGATGGGGTTTCACCATGTGGGCGAGCTGGTCTCAAACTCCGACCTCAGA 4924
QY 301 G 301
DB 4924 G 4924

```

RESULT 3
US-09-810-673A-3/c
: Sequence 3, Application US/09810673A
: Patent No. US20020064821A1
: GENERAL INFORMATION:
: APPLICANT: NERKULOV, Gennady et al.
: TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS
: TITLE OF INVENTION: AND USES THEREOF
: FILE REFERENCE: CL000723
: CURRENT APPLICATION NUMBER: US/09/810,673A
: CURRENT FILING DATE: 2001-03-19
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 26657
: TYPE: DNA
: ORGANISM: Human
US-09-810-673A-3

```

| Query Match | 67.8% | Score 204 | DB 9 | Length 26657 |
|-----------------------|----------------|---|--|--------------|
| Best Local Similarity | 87.4% | Pred. No. 5.4e-38 | | |
| Matches 222 | Conservative 1 | Mismatches 31 | Indels 0 | Gaps 0 |
| Qy | 44 | TATTAAGCTCTTTT | TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTAGATGAGATTGTGCTGTGTTG | 103 |
| Db | 445 | TTTGTAGATGAGATTGTGCTGTATATG | 386 | |
| Qy | 104 | CCCATGCTGAGTGGAGATGGCATGATCATATGCTCACTGCACCTCCACTCTGGGTTTA | 163 | |
| Db | 385 | CCCAAGGCTGGATGGCATGGACATGCAATCTCAGCTCAACAACCTCCGCTCCCGGGTTTCA | 326 | |
| Qy | 164 | AGCAAAAGCTGTGGCTCAGCCTCCGCGGATGATGGAGTTACAGGGGCCCAACACACACAAT | 223 | |
| Db | 325 | AGCATATGCTCCGCGCTCAGCCTCCCAAGATGCTGATTTACAGGGCACACACACACAGCTC | 266 | |
| Qy | 224 | CGGCTATATGTTGTATTTTATGTAGATGAGAGGGGTTTACATGTTGGCCAGGCTGGTCTC | 283 | |
| Db | 265 | CGGCTAATATTTTGTATTTTATGTAGAGAGGGGGTTTACCATGTTGGTTAGGCTGGTCTC | 206 | |
| Qy | 284 | AAACTCTGACCTC | 297 | |
| Db | 205 | AAACTCTGACCTC | 192 | |

```

1 RESULT 4
2 US-10-395-242-3/C
3 ; Sequence 3, Application US/10395242
4 ; Publication No. US20030165154A1
5 ;
6 ; GENERAL INFORMATION:
7 ;
8 ; APPLICANT: MERTIKOV, Gennady et al.
9 ;
10 ; TITLE OF INVENTION: ISOLATED HUMAN ION CHANNEL PROTEINS,
11 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN ION CHANNEL PROTEINS,
12 ; TITLE OF INVENTION: AND USES THEREOF
13 ;
14 ; FILE REFERENCE: C100723CON
15 ;
16 ; CURRENT APPLICATION NUMBER: US/10/395,242
17 ;
18 ; CURRENT FILING DATE: 2003-03-25
19 ;
20 ; NUMBER OF SEQ ID NOS: 7
21 ;
22 ; SOFTWARE: FastSeq for Windows Version 4.0
23 ;
24 ; SEQ ID NO 3
25 ;
26 ; LENGTH: 26657
27 ;
28 ; TYPE: DNA
29 ;
30 ; ORGANISM: Human
31 ;

```


106 C A T C T T G G A G T T G G A T T G C A T G A C T T C A C T T C C A C C T C C T G G T T C A A G 165

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2004, 17:19:44 ; Search time 2439.15 seconds

(Without alignments)
4342.094 Million cell updates/sec

Title: US-09-966-880A-9_COPY_5214_5514

Perfect score: 301
Sequence: 1 cctctctattcttcatagtc.....tcaactcctgcaccctcagag 301

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3757730 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA Main:*

1: /cgn2_6/ptodata/2/pna/PCRNUS_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US092A_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US092B_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US093A_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US093B_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US095D_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US097A_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US097B_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US097C_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US098A_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US098B_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US098C_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US098D_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US099D_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US099E_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US099F_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US100A_COMB.seq.*

44: /cgn2_6/ptodata/2/pna/US100B_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US101A_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US101B_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US102A_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US102B_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US103A_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US103B_COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US104A_COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US104B_COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US106_COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US107A_COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US107B_COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6000_COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6001_COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6002_COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6003_COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6004_COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6005_COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6007_COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6008_COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6009_COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6010_COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6011_COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6012_COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6013_COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6014_COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6015_COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6016_COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6017_COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6018_COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6019_COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6020_COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US6021_COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US6022_COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US6023A_COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US6023B_COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US6023C_COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US6024_COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US6025_COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US6026_COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US6027_COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US6028_COMB.seq.*
86: /cgn2_6/ptodata/2/pna/US6029_COMB.seq.*
87: /cgn2_6/ptodata/2/pna/US6030_COMB.seq.*
88: /cgn2_6/ptodata/2/pna/US6031_COMB.seq.*
89: /cgn2_6/ptodata/2/pna/US6032_COMB.seq.*
90: /cgn2_6/ptodata/2/pna/US6033_COMB.seq.*
91: /cgn2_6/ptodata/2/pna/US6034_COMB.seq.*
92: /cgn2_6/ptodata/2/pna/US6035_COMB.seq.*
93: /cgn2_6/ptodata/2/pna/US6036_COMB.seq.*
94: /cgn2_6/ptodata/2/pna/US6037_COMB.seq.*
95: /cgn2_6/ptodata/2/pna/US6038_COMB.seq.*
96: /cgn2_6/ptodata/2/pna/US6039_COMB.seq.*
97: /cgn2_6/ptodata/2/pna/US6040_COMB.seq.*
98: /cgn2_6/ptodata/2/pna/US6041_COMB.seq.*
99: /cgn2_6/ptodata/2/pna/US6042_COMB.seq.*
100: /cgn2_6/ptodata/2/pna/US6043_COMB.seq.*
101: /cgn2_6/ptodata/2/pna/US6044_COMB.seq.*
102: /cgn2_6/ptodata/2/pna/US6045_COMB.seq.*
103: /cgn2_6/ptodata/2/pna/US6046_COMB.seq.*
104: /cgn2_6/ptodata/2/pna/US6047_COMB.seq.*
105: /cgn2_6/ptodata/2/pna/US6048_COMB.seq.*
106: /cgn2_6/ptodata/2/pna/US6049_COMB.seq.*
107: /cgn2_6/ptodata/2/pna/US6050_COMB.seq.*
108: /cgn2_6/ptodata/2/pna/US6051_COMB.seq.*
109: /cgn2_6/ptodata/2/pna/US6052_COMB.seq.*
110: /cgn2_6/ptodata/2/pna/US6053_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 5514
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: intron
; LOCATION: (1)...(1031)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1032)...(1118)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (1119)...(5514)
US-09-966-880A-9

Query Match          99.9%; Score 300.6; DB 41; Length 5514;
Best Local Similarity 100.0%; Pred. No. 1,8e-25;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CTCTTCATATTTTCCATAGATATTCAAGTTTGACATATGATCAAGATATTACCTTCTTTT 60
Db      5214  CTCTTCATATTTTCCATAGATATTCAAGTTTGACATATGATCAAGATATTACCTTCTTTT 5273

QY      61  TTTTTTTTTTTTTTTTTTTGAGATGGAGTTTGCTCTTGTTGGCCATGCTGAGTGGAA 120
Db      5274  TTTTTTTTTTTTTTTTTTTGAGATGGAGTTTGCTCTTGTTGGCCATGCTGAGTGGAA 5333

QY      121  TGGCATGATCATAGCTCACTGCAACCTCCACTCTGGGTTCAAGCAAAAGCTGTGCGCTC 180
Db      5334  TGGCATGATCATAGCTCACTGCAACCTCCACTCTGGGTTCAAGCAAAAGCTGTGCGCTC 5393

QY      181  AGCCTCCCGGATGATGGGATTACAGGCGCCCAACCACTGCGCTAATGTTGTATT 240
Db      5394  AGCCTCCCGGATGATGGGATTACAGGCGCCCAACCACTGCGCTAATGTTGTATT 5453

QY      241  TTTAGTAGAGATGGGGTTTCAACCATGTTGGCCAGGCTGGTCAAACCTCGACCTAGA 300
Db      5454  TTTAGTAGAGATGGGGTTTCAACCATGTTGGCCAGGCTGGTCAAACCTCGACCTAGA 5513

QY      301  G 301
Db      5514  G 5514

```

| | | | |
|----|------|--|------|
| QY | 61 | TTTTTTTTTTTTTTTTTTTGGAGTGGAGTTTGGCTGTGTGCCCAAGCTGGAGTGGAA | 120 |
| Db | 4684 | TTTTTTTTTTTTTTTTTTTGGAGTGGAGTTTGGCTGTGTGCCCAAGCTGGAGTGGAA | 4744 |
| QY | 121 | TGGCATGAYCATAGCTCACTGCACCTCCACTCTGGATTCAAGCAAGCTGTGGCTC | 180 |
| Db | 4744 | TGGCATGACATAGCTCACTGCACCTCCACTCTGGATTCAAGCAAGCTGTGGCTC | 4803 |
| QY | 181 | AGCTCCCGGGGTAGATGGGATTACAGGGGCCACACACACACTGGGGCTAATGTGTGTANT | 240 |
| Db | 4804 | AGCTCCCGGGGTAGATGGGATTACAGGGGCCACACACACACTGGGGCTAATGTGTGTANT | 4863 |
| QY | 241 | TTTAGTAGAGATGGGGTTTCAACATGTTGGCCAGGCTGGTCTCAAACTCTGACCTCAGA | 300 |
| Db | 4864 | TTTAGTAGAGATGGGGTTTCAACATGTTGGCCAGGCTGGTCTCAAACTCTGACCTCAGA | 4922 |
| QY | 301 | G 301 | |
| Db | 4924 | G 4924 | |

```

RESULT 3
US-09-966-880A-35
; Sequence 35, Application US/09966880A
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; CURRENT FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 11204
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-966-880A-35

```

| | | | | |
|---------------------------|--------|--------------------|-----------|---------------|
| Query Match | 99.9%; | Score 300.6; | DB 41; | Length 11204; |
| Best Local Similarity | 99.7%; | Pred. No. 1.5e-25; | | |
| Matches 300; Conservative | 1; | Mismatches 0; | Indels 0; | Gaps 0; |

| QY | 1 | CTCTCTCAATTTTTCACATAGTATCAAGTTTGACATATCAAGTTTACTCTTTCTTTT | 60 |
|----|------|--|------|
| Db | 4684 | CTCTCTCAATTTTTCACATAGTATCAAGTTTGACATATCAAGTTTACTCTTTCTTTT | 4683 |
| QY | 61 | TTTTTTTTTTTTTTTTTTTGAGATGAGCTTTGGTCTTGTGTCCCATGCTGGAAGTGA | 120 |
| Db | 4684 | TTTTTTTTTTTTTTTTTTTGAGAGGAGTTTGGTCTTGTGTCCCATGCTGGAAGTGA | 4743 |
| QY | 121 | TGGCATGACATAGTCACTCACTCACTCTCTGGGTTCAAGCAAGCTGTGCGCTC | 180 |
| Db | 4744 | TGGCATGACATAGTCACTCACTCACTCTCTGGGTTCAAGCAAGCTGTGCGCTC | 4803 |
| QY | 181 | AGCTCTCCGGGTAGATGGGATTTACAGGCGCCACACACACACTCGGCTAATGTTGTATT | 240 |
| Db | 4804 | AGCTCTCCGGGTAGATGGGATTTACAGGCGCCACACACACTCGGCTAATGTTGTATT | 4863 |
| QY | 241 | TTTAGTAGAGATGGGGTTTCAACATAGTTGGCCAGGCTGATCTCAAACTCTGACCTCAGA | 300 |
| Db | 4864 | TTTAGTAGAGATGGGGTTTCAACATAGTTGGCCAGGCTGATCTCAAACTCTGACCTCAGA | 4923 |
| QY | 301 | G 301 | |

Db 4924 G 4924

RESULT 4
US-09-949-016-13942

```

; sequence 13942, Application US/09949016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

```

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE NUMBER: 001303

```

; FIDEREERENCE: CUV0130,
;
; CURRENT APPLICATION NUMBER: US/09/949,016
;
; CURRENT FILING DATE: 2000-04-14

```

1 PRIOR APPLICATION NUMBER: 60/241,755
2
3
4
5 PRIOR FILING DATE: 2000-10-20
6
7
8
9 PRIOR APPLICATION NUMBER: 60/237,768
10

PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498

```

; PRIOR FILING DATE: 2000-09-08
;
; NUMBER OF SEQ ID NOS: 207012
;
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 13942
; LENGTH: 14684
;

```

TYPE: DNA
ORGANISM: Human
US-09-949-016-13942

| | | | | |
|---------------------------|-------|--------------------|-----------|---------------|
| Query Match | 99.9% | Score 300.6; | DB 39; | Length 14684; |
| Best Local Similarity | 99.7% | Pred. No. 1.5e-25; | | |
| Matches 300; Conservative | 1; | Mismatches 0; | Indels 0; | Gaps 0 |

Dy 1 CTCCTCAATTTTCATAGATTCAAGTTGACAAATGATCAAGTATTACTCTTTCTTTTT 60
| | | | |
Db 6180 CTCTCTAATTTTTCATAGATTCAAGTTGACAAATGATCAAGTATTACTCTTTCTTTTT 6330

61 TTTTCTTTTTTTTTTTGAGATGAGCTTTGGCTGTGTGCCCATGCTGGAGTGGAA 120

D6 6240 TTTTCTTTTTTTTTTTTTTGAGATGGAGTTTTCGTCCTGTGCCCATGCCTGGAGTGGAA 629

D7 121 TGGCATGAYCATAGCTCACTGCACCTCCACTCCTGGGTTCAAGCAAAGCTGTGGCTC 180

| | | | |
|----|------|---|------|
| Db | 6300 | TGGCATGACCATAGCTCACTGCGAACCTCCAGCTCTGCGCTC | 6350 |
| Oy | 181 | AGCCTCCCGGGTAGATGGCATTTACAGGCGGCCACCAACAACCTCGGCTAATGTTTGTAAT | 240 |

6360 AGCTCCCGGTGATGGATTACAGGCGCCACCACTGGCTAATGTTGATT 6411

241 TTAGTAGAGATGGGGTTTACCATGTTGGCCAGGCTGTTCTCAACTCTGACCTCAGA 6470
 Db 6420

| | |
|----|-------------|
| QY | 301 G 301 |
| | — |
| Db | 6480 G 6480 |

RESULT 5
US-60-212-664-16

```

; Sequence 16, Application US/60212664
; GENERAL INFORMATION:
; APPLICANT: Ladunga, Steve
; INVENTOR:

```

```

: APPLICANT: Spier, gene
: APPLICANT: Greenberg, Simon
: APPLICANT: Radkin, Steven
: APPLICANT: Wang, Yu

```

TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEIN
 TITLE OF INVENTION: USES THEREOF

```

: FILE REFERENCE: CU000687
:
: CURRENT APPLICATION NUMBER: US/60/212,664
:
: CURRENT FILING DATE: 2000-06-20
:
: NUMBER OF SEQ ID NOS: 636

```

```
SOFTWARE: FastSeq for Windows Version 4.0
```

SEQ ID NO 16
LENGTH: 155313
TYPE: DNA
ORGANISM: HUMAN
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)-(155313)
OTHER INFORMATION: n = A,T,C or G
US-60-212-664-16

Query Match 99.9%; Score 300.6; DB 77; Length 155313;
Best Local Similarity 99.7%; Pred. No. 9e-26;
Matches 300; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCTCTATTTTCCATAGTATTCAGTTTGACATGATCAAGTATTACTCTTTCTTTT 60
DB 123720 CTCTCTATTTTCCATAGTATTCAGTTTGACATGATCAAGTATTACTCTTTCTTTT 123779
QY 61 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 120
DB 123780 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 123839
QY 121 TGGCATGAYCATAGCTCACTGCACTCCACTCTCTGAGTTTCAAGCAAAAGCTGCTCCCTC 180
DB 123840 TGGCATGACCATAGCTCACTGCACTCCACTCTCTGAGTTTCAAGCAAAAGCTGCTCCCTC 123899
QY 181 AGCTCCCGGGTAGATGAGATTACAGGCGCCACACACACACTGCGCTAATGTTGTATT 240
DB 123900 AGCTCCCGGGTAGATGAGATTACAGGCGCCACACACACACTGCGCTAATGTTGTATT 123959
QY 241 TTTAGTAGAGATGGGGTTTCACTGATGTTGGCCAGGCTGCTCAAACTCTGACCTCAGA 300
DB 123960 TTTAGTAGAGATGGGGTTTCACTGATGTTGGCCAGGCTGCTCAAACTCTGACCTCAGA 124019
QY 301 G 301
DB 124020 G 124020

RESULT 6
US-60-465-241-52403
Sequence 52403, Application US/60465241
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001468
CURRENT APPLICATION NUMBER: US/60/465,241
CURRENT FILING DATE: 2003-04-23
NUMBER OF SEQ ID NOS: 258418
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 52403
LENGTH: 22683
TYPE: DNA
ORGANISM: Homo sapiens
US-60-465-241-52403

Query Match 99.6%; Score 299.8; DB 103; Length 22683;
Best Local Similarity 99.0%; Pred. No. 1.6e-25;
Matches 298; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCTCTATTTTCCATAGTATTCAGTTTGACATGATCAAGTATTACTCTTTCTTTT 60
DB 10179 CTCTCTATTTTCCATAGTATTCAGTTTGACATGATCAAGTATTACTCTTTCTTTT 10238
QY 61 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 120
DB 10238 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 10298
QY 121 TGGCATGAYCATAGCTCACTGCACTCCACTCTCTGAGTTTCAAGCAAAAGCTGCTCCCTC 180
DB 10299 TGGCATGACCATAGCTCACTGCACTCCACTCTCTGAGTTTCAAGCAAAAGCTGCTCCCTC 10358

QY 181 AGCTCCCGGGTAGATGAGATTACAGGCGCCACACACACACTGCGCTAATGTTGTATT 240
DB 10359 AGCTCCCGGGTAGATGAGATTACAGGCGCCACACACACACTGCGCTAATGTTGTATT 10418
QY 241 TTTAGTAGAGATGGGGTTTCACTGATGTTGGCCAGGCTGCTCAAACTCTGACCTCAGA 300
DB 10419 TTTAGTAGAGATGGGGTTTCACTGATGTTGGCCAGGCTGCTCAAACTCTGACCTCAGA 10478
QY 301 G 301
DB 10479 G 10479

RESULT 7
US-60-470-166-8933
Sequence 8933, Application US/60470166
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001464
CURRENT APPLICATION NUMBER: US/60/470,166
CURRENT FILING DATE: 2003-05-09
NUMBER OF SEQ ID NOS: 68617
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8933
LENGTH: 22683
TYPE: DNA
ORGANISM: Homo sapiens
US-60-470-166-8933

Query Match 99.6%; Score 299.8; DB 104; Length 22683;
Best Local Similarity 99.0%; Pred. No. 1.6e-25;
Matches 298; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCTCTATTTTCCATAGTATTCAGTTTGACATGATCAAGTATTACTCTTTCTTTT 60
DB 10179 CTCTCTATTTTCCATAGTATTCAGTTTGACATGATCAAGTATTACTCTTTCTTTT 10238
QY 61 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 120
DB 10238 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 10298
QY 121 TGGCATGAYCATAGCTCACTGCACTCCACTCTCTGAGTTTCAAGCAAAAGCTGCTCCCTC 180
DB 10299 TGGCATGACCATAGCTCACTGCACTCCACTCTCTGAGTTTCAAGCAAAAGCTGCTCCCTC 10358
QY 181 AGCTCCCGGGTAGATGAGATTACAGGCGCCACACACACACTGCGCTAATGTTGTATT 240
DB 10359 AGCTCCCGGGTAGATGAGATTACAGGCGCCACACACACACTGCGCTAATGTTGTATT 10418
QY 241 TTTAGTAGAGATGGGGTTTCACTGATGTTGGCCAGGCTGCTCAAACTCTGACCTCAGA 300
DB 10419 TTTAGTAGAGATGGGGTTTCACTGATGTTGGCCAGGCTGCTCAAACTCTGACCTCAGA 10478
QY 301 G 301
DB 10479 G 10479

RESULT 8
US-09-947-911-286
Sequence 286, Application US/09947911
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig
TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED ON
FILE REFERENCE: C1001300
CURRENT APPLICATION NUMBER: US/09/947,911
CURRENT FILING DATE: 2001-09-07
NUMBER OF SEQ ID NOS: 368

```

? SEQ ID NO 286
? LENGTH: 3037227
? TYPE: DNA
? ORGANISM: HUMAN
? FEATURE:
? NAME/KEY: m1c_feature
? LOCATION: (1,...,3037227)
? OTHER INFORMATION: n = A,T,C or G
? JS-09-947-911-286

```

| | | | | |
|---------------------------|--------|--------------------|-----------|-----------------|
| Query Match | 69.4%; | Score 208.8; | DB 39; | length 3037227; |
| Best Local Similarity | 88.6%; | Pred. No. 1.1e-15; | | |
| Matches 225; Conservative | 1; | Mismatches 28; | Indels 0; | Gaps 0; |

| | | | |
|----|---------|---|---------|
| Gy | 46 | TTCATCTTCTTTTCTTTTTTTTTTTTTTTTGAGATGGAGTTTGGTCTGTGACC | 105 |
| Dd | 1512582 | TTCTCTTCCCTTTTTTTTTTTTTTTTGAGATGGAGTTTGCCTCTGTGCC | 1512641 |

0y 10c CATCTGGAGTGCATGCGATGATCATCTGCAACCTCCACCTCTGGATTCAAG 165
1512642 CAGCTGGAGTGCATGCGGTGATCTCACTGCAACCTCCACCTCTGGATTCAAG 1512700

| Qy | 166 | CGAAGCTGTGCGCTCAGCCTCCCGGGTATGAGGATTACAGCGGCCACACACACACACTCG | 225 |
|----|---------|--|---------|
| Db | 1512702 | CGATTCTCCTGCCCTCAGCCTCCTAGTAGTGGATTACAGTGGCTGCTGCCACACACCCA | 1512761 |

| | | | |
|----|---------|--|---------|
| OY | 22 | GCATATCTTGTATTTTGTAGAGATGGGGTTTCCACATGTTGGCCAGGCTGATCCAA | 285 |
| Db | 1512762 | GCTATCTTTGTATTTTGTAGAGACAGGGTTTCACCATGTTGGCCAGGCTGCTCCGA | 1512821 |

| | | | |
|-----------|----------------|----------------|---------|
| OY | 286 | ACTCTGACCTCAG | 299 |
| | | | |
| | | | |
| | | | |
| Db | 1512822 | ACTCCTGAGCTCAG | 1512835 |

RESULT 9
PCT-US01-21148-3/C
Sequence 3 Application PC/TTTS0121148

```

; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al.
; TITLE OF INVENTION: ISOLATED TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NITRIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: TRANSPORTER PROTEININCS

```

```

1  TITLE OF INVENTION:  AND USES THEREOF
2
3  FILE REFERENCE:  CL000723PCT
4
5  CURRENT APPLICATION NUMBER:  PCT/US01/21148
6
7  CURRENT FILING DATE:  2001-07-03

```

? PRIOR APPLICATION NUMBER: 60/216,340
 ? PRIOR FILING DATE: 2000-07-05
 ? PRIOR APPLICATION NUMBER: 09/810,673
 ? PRIOR FILING DATE: 2001-03-19

```

; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 26657

```

```

;             IIR: DNA
; ORGANISM: Human
PCR-US01-21148-3

```

| | | | |
|---------------------------|--------|------------------|-------------------|
| Query Match | 87.48; | Score 204; LB 1; | Denglin 2003; |
| Best Local Similarity | 87.48; | Pred. No. 1e-14; | |
| Matches 222; Conservative | 1; | Mismatches 31; | Indels 0; Gaps 0; |

[illegible]

385 CCCAGGGCTGAGTGCATGSCACGATCTCAGCTCACCAACCTCCGCTCCCCGGATTCA 326

325 AGCAATGCTCCCGCCTCAGCCTCCCAAGTAGCTGAGATTACAGGCGCACACCAACGCC 286

| Accession | Sequence | Position |
|-----------|---|----------|
| Oy | CGGCTATGTTTATTTTATAGAGATGGGGTTTCACTGTTGGCCAGGCTGCTC | 283 |
| Oy | 22 CGGCTATGTTTATTTTATAGAGATGGGGTTTCACTGTTGGCCAGGCTGCTC | 283 |
| Db | CGGCTATTTTATTTTATTTTATAGAGACGGGGTTTCACTGTTGGTTAGGCTGCTC | 206 |
| Db | 265 CGGCTATTTTATTTTATTTTATAGAGACGGGGTTTCACTGTTGGTTAGGCTGCTC | 206 |
| Oy | 284 AAACCTCTGACCTTC | 297 |
| Oy | 205 AAACCTCTGACCTTC | 192 |
| Db | | |

RESULT 10
US-09-810-673A-3/C
Sequence 3 Amplification ITS/09810673A

```

: GENERAL INFORMATION:
:
: APPLICANT: MERKLOV, Gennady et al.
:
: TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
:
: TITLE OF INVENTION: NICKETIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS

```

; TITLE OF INVENTION: AND USES THEREOF
 ; FILE REFERENCE: CL000723
 ; CURRENT APPLICATION NUMBER: US/09/810,673A
 ; CURRENT FILING DATE: 2001-03-19

```

; NUMBER OF SEQ ID NOS: 7
;
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 3
;
; LENGTH: 26657
;

```

```

!      TYPE: DNA
!      ORGANISM: Human
US-09-810-673A-3

```

| | | | | |
|---------------------------|-------|-----------------|----------|--------------|
| Query match | 67.8% | Score 204 | DB 337 | Length 20037 |
| Best local Similarity | 87.4% | Pred. No. 1e-14 | | |
| Matches 222, Conservative | 1 | Mismatches 31 | Indels 0 | Gaps 0 |

[illegible]

Db 385 CCCAGGCTGGAGTGCATGGCAGATCTACGTCACACCAACCTCCGGCTCCGGGGTTCA 328

DB 325 AGCAATGCTCCGCGCTCAGCTCCCAAGTAGCTGAGATTACAGGCGCACCAACCGCC 266

Db 265 CGGCTAATTTTGTATTTAGTAGACGGGGTTTCACCAAGTTGGCTTAGGCTGCTTC 206

205 AAACCTCCTGACCTC 192

RESULT 11
US-10-395-242-3/C
; Sequence 3, Application US/10395242

APPLICANT: MERKULOV, Gennady et al.
TITLE OF INVENTION: ISOLATED HUMAN ION CHANNEL PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN ION CHANNEL PROTEINS

```

FILE REFERENCE: CL000723CON
CURRENT APPLICATION NUMBER: US/10/395,242
CURRENT FILING DATE: 2003-03-25

```

```

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 26657
;

```

! ORGANISM: Human
US-10-395-242-3

| | | | |
|----|---------|---|----------|
| OY | 44 | TATTAACCTTCTCTTTTTTTTTTTTTTTTTTTTTTGAGANGGAGTTTGTCGTGTG | 103 |
| Db | 2278243 | TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGANGAAGTTTGCCTCTATTG | 22781649 |
| OY | 104 | CCCATGCTGAGATGGAGATGGCATTGAYCAATAGCTACTGAACTTCACCTCTCTGGCTTCA | 163 |
| Db | 2278183 | CCCAGGCTGGAGATGGATGGACAAGATCTCAGCTCCCAACCTCTCCCGGCTTCA | 22781244 |
| OY | 164 | AGCAAAGCTGACGCTCAGCCTCCCGGGTAGTGGATTACAGGCGGCCACACACACT | 223 |
| Db | 2278123 | AGCAATGCTCCCGCCTCAGCCTCCCAAGTAGCTGAGATTACGGGCGCACACACAGCC | 22780649 |
| OY | 224 | CGGCTAATGTTGTGATTTTTAGTAGAGATGGGGTTTACCAGTTTGCCAGGCTGGTCTC | 283 |
| Db | 2278063 | CGGCTAATTTTTGTGATTTTTAGTAGAGACGGGGTTTCAACATGTTGGTTAGGCTGGTCTC | 22780049 |
| OY | 284 | AAACTCTGACCTC 297 | |
| Db | 2278003 | AAACTCTGACCTC 22777990 | |

```

RESULT 15
Sequence 43: Application US/60206600
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS OF THE
TITLE OF INVENTION: SER/THR AND TYR FAMILY OF KINASES, NUCLEIC ACID MOLECULES
TITLE OF INVENTION: ENCODING THESE HUMAN KINASE PROTEINS, AND USES THEREOF
PUB REFERENCE: CLO00627
CURRENT APPLICATION NUMBER: US/60/206,600
CURRENT FILING DATE: 2000-05-31
NUMBER OF SEQ ID NOS: 172
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43
LENGTH: 32768
TYPE: DNA
ORGANISM: HUMAN
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)..(32768)
OTHER INFORMATION: n = A,T,C or G
US-60-206-600-43

```

| | | | | |
|-----------------------|----------------|------------------|----------|--------------|
| Query Match | 67.4% | Score 202.8 | DB 76 | Length 32768 |
| Best Local Similarity | 85.2% | Pred No. 1.3e-14 | | |
| Matches 225 | Conservative 1 | Mismatches 38 | Indels 0 | Gaps 0 |

| | | | | |
|----|--|------|---|------|
| QY | | 36 | TGATCAAACTAATTACCTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTGTAGATGGAGTTTG | 95 |
| Dd | | 8722 | TGACCATAATTCCTTTTTTTTTTTCTTTCTTTTTTTTTTTTTGAATGGAGTTTGC | 866 |
| QY | | 96 | TCTTGTGCCCATGCTGSAGTGAATGGCAGTAATAGACTCACTGACACTCCACTCC | 155 |
| Dd | | 8662 | TCTTGTTCACACAGCGCTGSAGTGCATAGGCACAGATTCGGCCCACTGCACAACCTCCCCTCC | 8603 |
| QY | | 156 | TGGGTTCAAGCAAAGCTGTGGCTTAGGCTCCCCGGGTAGATGGGATTTACAGGCGGCCAC | 215 |
| Dd | | 8602 | TGGGTTCAAGCAATTCCTCTGCTTAGGCTCCCGAGTACGTGGGATTACAGCGCTCAC | 854 |
| QY | | 216 | ACCACACTCGGCTAAATGTTGATTTTGTATTTAGTAGAGATGGGTTTCAACATGTTGGACA | 275 |
| Dd | | 8542 | ACCACACCAAGCTAATTTTTTGTATTTTATTTAGTAGAGAAGGGGTTTCACTGTGTTGGCTAG | 8483 |
| QY | | 276 | CTGGTCTCAAACTCCTGACCTCAG | 299 |
| Dd | | 8482 | CTGGTCTCAAACTCCTGACCTCAG | 8459 |

```
Search completed: March 12, 2004, 09:13:14
Job time : 2466.15 secs
```


TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences

218 CACACTCGGCTATGTTTGTATTTTACAGAGATGGGTTTCACCATGTTGGCAGGT 277

